11-7-2011

Angular Distribution of Z0 Bosons in Z+Jet Events at sqrt(S) = 7 TeV

Luis Lebolo
Florida International University, luis.lebolo@fiu.edu

DOI: 10.25148/etd.FI11120511

Recommended Citation

Lebolo, Luis, "Angular Distribution of Z0 Bosons in Z+Jet Events at sqrt(S) = 7 TeV" (2011). *FIU Electronic Theses and Dissertations*. 492.
https://digitalcommons.fiu.edu/etd/492

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

ANGULAR DISTRIBUTION OF Z^0 BOSONS IN Z+JET EVENTS AT $\sqrt{S} = 7$ TEV

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

by

Luis Lebolo

2011

To: Dean Kenneth Furton
College of Arts and Sciences

This dissertation, written by Luis Lebolo, and entitled Angular Distribution of Z^0 Bosons in Z+Jet Events at $\sqrt{S} = 7$ TeV, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

Stephan Linn

Jorge Rodriguez

Joerg Reinhold

Masoud Sadjadi

Pete Markowitz, Major Professor

Date of Defense: November 7, 2011

The dissertation of Luis Lebolo is approved.

Dean Kenneth Furton
College of Arts and Sciences

Dean Lakshmi N. Reddi
University Graduate School

Florida International University, 2011

DEDICATION

I dedicate this work to the Giants in my life. Without their shoulders to stand on...

I would be lost in the crowd.

> If I have seen further it is by standing on the shoulders of giants
>
> –Sir Isaac Newton, 1676

ACKNOWLEDGMENTS

I would like to express my appreciation for the following Giants:

Alan “Goober” Flaten – Excuse The Flagrant Semicolon Usage

Andrew “Colúmbian” Kubik – Forever HCAL

Brian “B′” Beckford – Still Comin’ Strong

Cristobal “Penthus” Ceron – Ordained To Be Forever Unaccompanied In Obscurity

Daniel “Delish” Christoph – The Ring-Bearer

Edward “Behbeh” Laird – Don’t Ted Out, Man

Ellie “HCAL” Twedt – Undercover Chihuahua Lover

Fareed “Black Ice” Khan – Double Quilted Shoulder Upper

Gerald “Forever” Lebolo – How’s About We Get That Tattoo

German “Empanadero” Martinez – El Hermano Mayor Cuida Al Cachorro

Javier “Wudope” Beteta – The First Of The Last To Get Married

Philip “Lookin’ Good” Lawson – Mexican Omelette Time!

Priscilla “Cybele” Pamela – Creating Triangles Out Of Straight Lines

Private “No. 65” Ham – You Travel The World In Our Stead

Seth “Maybelline” Manthey – Hyena > Cheetah

Vanessa “Hey Luis!” Werner – Bruiser Of Sidewalks

For my beloved liChi, Alicia Redondo; años de estudiar el universo no me han preparado para la alegría, el entusiasmo y el amor en que me has envuelto - te amo.

For my mother, Yomaira Lebolo, and my father, Jairo Lebolo; sus innumerables sacrificios me han convertido en el hombre que soy hoy en dia.

For my brother, Jayder Lebolo, who has done the work of three sons. If the above are my Giants, then you are my Atlas - you carried my world on your shoulders.

For my nieces, Marilyn and Rinoa Lebolo, and my nephew, Jayden Lebolo; may this family bless you as it has surely blessed me.

For my grandmother, Santos Piocuda; a timeless smile. For my grandfather, Esteban Piocuda; a timeless sage.

For one of the most admirable individuals I have ever met, Vashti Sawtelle; your relentless “GO TEAM!” attitude will inspire and motivate me for years to come.

For mentor from the past, Dr. Brian Raue; thanks for allowing me to whine, for reminding me that “it’s just math,” and for always being blunt and genuine. It could not have been a defense without your (much appreciated) criticism.

For committee members, Dr. Joerg Reinhold and Dr. Masoud Sadjadi; thank you for taking a sincere interest in my success. For the only other particle physicist to come out of Carol City, Dr. Jorge Rodriquez; in retrospect, bumping heads with you has made this process worthwhile.

And for Dr. Stephan Linn and Dr. Pete Markowitz; together you have played the roles of advisor, father, brother, and friend, and have given me the greatest gift possible - a sound and logical (if somewhat frayed) mind.

Although it may seem like boilerplate, I would like to sincerely congratulate my colleagues in the CERN accelerator departments for the excellent performance of the LHC machine. I also thank the technical and administrative staff at CERN and other CMS institutes with support from: FMSR (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES (Croatia); RPF (Cyprus); Academy of Sciences and NICPB (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); OTKA and NKTH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); NRF and WCU (Korea); LAS (Lithuania); CINVESTAV, CONACYT, SEP, and UASLP-FAI (Mexico); MSI (New Zealand); PAEC (Pakistan); SCSR (Poland); FCT (Portugal); JINR (Armenia, Belarus, Georgia, Ukraine,

Uzbekistan); MST, MAE and RFBR (Russia); MSTD (Serbia); MICINN and CPAN (Spain); Swiss Funding Agencies (Switzerland); NSC (Taipei); TUBITAK and TAEK (Turkey); STFC (United Kingdom); DOE and NSF (USA).

ABSTRACT OF THE DISSERTATION

ANGULAR DISTRIBUTION OF Z^0 BOSONS IN Z+JET EVENTS AT $\sqrt{S} = 7$ TEV

by

Luis Lebolo

Florida International University, 2011

Miami, Florida

Professor Pete Markowitz, Major Professor

For the first time, the Z^0 boson angular distribution in the center-of-momentum frame is measured in proton-proton collisions at $\sqrt{s} = 7$ TeV at the CERN LHC. The data sample, recorded with the CMS detector, corresponds to an integrated luminosity of approximately 36 pb^{-1}. Events in which there is a Z^0 and at least one jet, with a jet transverse momentum threshold of 20 GeV and absolute jet rapidity less than 2.4, are selected for the analysis. Only the Z^0's muon decay channel is studied. Within experimental and theoretical uncertainties, the measured angular distribution is in agreement with next-to-leading order perturbative QCD predictions.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

ATLAS	A Toroidal LHC Apparatus
CDF	Collider Detector at Fermilab
CERN	European Organization for Nuclear Research
CM	center of momentum
CMS	Compact Muon Solenoid
CSC	cathode strip chamber
D0	DZero Experiment
DT	drift tube
ECAL	electromagnetic calorimeter
GeV	giga-electron volt
HCAL	hadron calorimeter
HLT	high-level trigger
IC	iterative cone
IP	interaction point
IRC	infrared and collinear
JES	jet energy scale
L1	Level-1
LHC	Large Hadron Collider
LO	leading-order
MC	Monte Carlo
ME	matrix element
MeV	Mega-electron volt
NDOF	number of degrees of freedom
NLO	next-to-leading-order
NNLO	next-to-next-to-leading-order

PAT	physics analysis toolkit
pb	picobarn
PDF	parton distribution function
PF	particle-flow
pp	proton-proton
pQCD	perturbative quantum chromodynamics
PS	parton showering
PV	primary vertex
QCD	quantum chromodynamics
QED	quantum electrodynamics
RPC	resistive plate chamber
SM	Standard Model
TeV	tera-electron volt

CHAPTER 1

INTRODUCTION

Particle physics provides a description of the universe at the most fundamental level by studying the basic building blocks of matter and how they interact. The culmination of these studies resulted in a theory called the Standard Model (SM). Much of the success attributed to the SM was obtained through the use of particle accelerators, which allowed physicists to probe unknown territories. Today's particle accelerators, like the Large Hadron Collider (LHC) at CERN, enable physicists to look even deeper into the subatomic scale and cross-check SM predictions.

The purpose of this dissertation is to measure the angular distribution of the Z^0 boson, in association with jets, in the parton-parton center of momentum (CM) frame. These events are produced at the LHC via proton-proton collisions (at $\sqrt{S} = 7$ TeV) and recorded using the Compact Muon Solenoid (CMS) detector. In general, events with a Z^0 accompanied by jets can be used for detector calibration, Monte Carlo generator tuning, and to model and constrain the background to many new physics searches (e.g., certain types of Higgs decays). In particular, the relative Z^0 angular distribution is measured because it can be decoupled from the parton distribution functions in the invariant cross section. This allows us to compare the measurement, with minimal complications, to perturbative quantum chromodynamics (pQCD) at the new energy regime of the LHC; a deviation from pQCD may signify weaknesses in the SM. To our knowledge, this analysis is the first measurement of the Z^0 angular distribution.

Chapter 1 introduces the SM, giving a brief overview of the theory required to understand the analysis. Chapter 2 gives an overview of the accelerator and detector and how they are used to analyze the collision data. Chapter 3 describes the reconstruction algorithms utilized to create particle representations out of raw data.

Chapter 4 explains some of the techniques used to simulate quantum chromodynamics. Chapter 5 lists the selections applied in order to identify candidate particles. Chapter 6 explains the kinematics of the CM frame, where the angular distribution is measured. Chapter 7 discusses the corrections applied to the angular distribution, as well as the sources of uncertainty. Chapter 8 consists of the final results and conclusions. The appendices contain unfolding cross-checks, efficiency studies, differences in jet-handling algorithms, and the source code of particle representations.

1.1. The Standard Model

The SM is the most comprehensive and widely-accepted particle physics theory available [1]. It consists of twelve fundamental particles which, along with their respective antiparticles, are the constituents of matter. The SM also describes the strong, weak, and electromagnetic particle interactions. The gravitational force is not incorporated within the SM; it is the weakest of the four fundamental forces and, though it must be taken into account over large scales, it is negligible in individual particle interactions. The only predicted SM particle not yet observed experimentally is the hypothetical Higgs boson, devised to explain spontaneous symmetry breaking.

1.1.1. Quarks and Leptons

The SM describes two types of spin-$\frac{1}{2}$ particles that make up matter: quarks and leptons (collectively known as fermions). The respective antiparticles of all twelve fermions also exist and carry opposite quantum numbers. While leptons do not interact via the strong force, quarks combine to form other particles such as protons and neutrons (these composite particles are called hadrons). Fermions are categorized into three generations (the three leftmost columns of Fig. 1.1). For higher generations, the lifetime of the particle generally decreases. They decay to particles of the lowest generation, which explains why everyday matter is comprised of first generation particles only (i.e., up quarks, down quarks, and electrons). In fact, all of the atoms on the periodic table of elements are made up of protons (i.e., two up quarks and one down quark), neutrons (i.e., two down quarks and one up quark), and electrons; all first generation fermions.

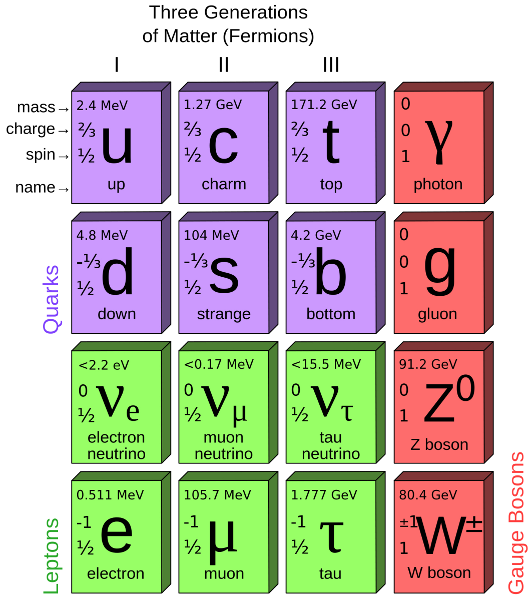


Figure 1.1: SM table of fermions and gauge bosons. Also shown (from top to bottom) is the mass, charge (in units of e), and spin of each particle. From left to right, each column of fermions corresponds to the first, second or third generation.

1.1.2. Forces

In the SM, every particle interaction is associated with a spin-1 mediator particle, i.e., particles that interact via a force do so through the exchange of gauge bosons (the rightmost column of Fig. 1.1). In order of decreasing force strength, the strong force is mediated by the gluon, the electromagnetic force is mediated by the photon, and the weak force is mediated by vector bosons ($W^{\pm}$ and Z^{0}). The electromagnetic and weak forces are 10^{2} and 10^{7} times weaker than the strong force, respectively. Each interaction also has an associated charge; the electromagnetic force is associated with the familiar electric charge, the weak force with flavor, and the strong force with color. While there is a hypothetical graviton for gravity, it is not discussed here. A summary of the interaction between particles is shown in Fig 1.2.

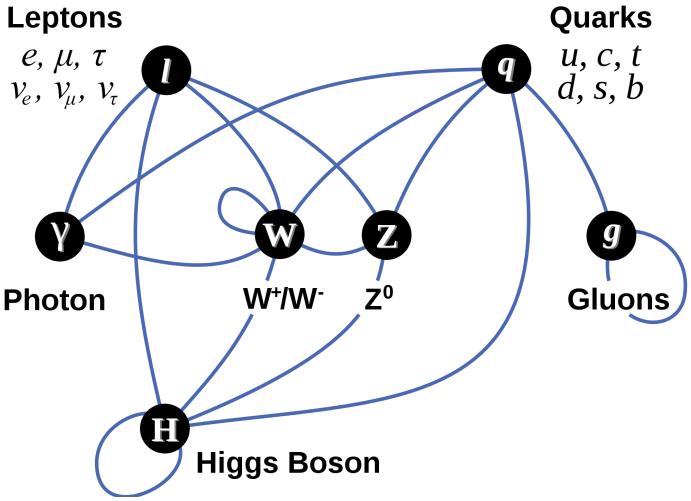


Figure 1.2: A diagram summarizing the SM interactions between particles. Vertices (circles) represent types of particles and edges (blue lines) connecting them represent interactions that can take place. The top row of vertices (fermions) are the matter particles, the second row of vertices (gauge bosons) are the force mediating particles, and the bottom row is the Higgs boson (not discussed). The Higgs interactions in this diagram represent how particles obtain mass.

Quantum Electrodynamics

Electromagnetism is described by a relativistic quantum field theory known as quantum electrodynamics (QED). Although the analysis tests quantum chromodynamics (QCD) predictions, important analogs can be made via the study of quantum electrodynamics. Indeed, much of the theoretical challenge of pQCD has been overcome by extending the "tried and true" methods of quantum electrodynamics. A very brief overview of QED can be obtained by examining the coupling strength of the electromagnetic force,

$$\alpha(Q^2) = \frac{\alpha(\Lambda^2)}{1 - \frac{\alpha(\Lambda^2)}{3\pi}\ln(\frac{Q^2}{\Lambda^2})}, \tag{1.1}$$

where Q is the momentum transfer and $\alpha(\Lambda^2)$ is the fine structure constant as a function of Λ, the energy scale. Conventionally, the fine structure constant is evaluated as 1/137 at large distances (i.e., low energy scales). Note that as the momentum transfer of a particle collision increases so does the coupling strength.

Quantum Chromodynamics

Interactions involving the strong force are also described by a relativistic quantum field theory called QCD. Unlike QED (which has a single, neutral carrier), QCD allows for eight gluons that each carry a different combination of color charge. This causes QCD to be significantly different from QED since gluons can interact with other gluons. The coupling strength of the strong force can be expressed as [2]

$$\alpha_s(Q^2) = \frac{\alpha_s(\Lambda^2)}{1 + \frac{\alpha_s(\Lambda^2)}{12\pi}(33 - 2n_f)\ln(\frac{Q^2}{\Lambda^2})}, \tag{1.2}$$

where n_f is the number of active fermion generations. Note that as the momentum transfer of a particle collision increases the coupling strength decreases; this has profound implications for quarks. At small distance scales, the strength of the strong force becomes small (referred to as *asymptotic freedom*); i.e., at high momentum transfers, the strong interactions are simple and calculable via perturbation theory. Conversely, as the distance increases the force gets stronger, which ultimately restrains quarks within hadrons (a phenomenon called *confinement*).

In fact, only color-neutral combinations of quarks are ever observed in nature. One such combination is a quark-antiquark pair; the antiquark carries the opposite color so that the composite particle is color-neutral. Another combination is that of three quarks, where each quark carries a different color charge, such that the composite particle is again color-neutral. However, quarks are not the only constituents of a hadron; contained within the hadron are the gluons mediating the interactions between the quarks. These gluons can momentarily split to create short-lived, virtual quark-antiquark pairs (called sea quarks). The quarks or gluons within a hadron are called partons.

Suppose that in a collision one can free a parton from its hadronic confines. As the parton starts to move away from the rest of the hadron, the strength of the strong force between the two parts intensifies and increases the probability for QCD radiation. Thus, the parton will emit gluons, which will radiate quark-antiquark pairs, resulting in the emission of more gluons, etc. (this process is called *fragmentation*). Each new parton is essentially collinear with its parent. The forces involved quickly cause the collection of partons to be rearranged into color-neutral combinations. What emerges far enough from the collision point to be detected is a spray or "jet" of hadrons; hence one can never detect a free parton. The transformation of a quark/gluon into hadrons is called *hadronization*.

1.2. Collisions

In a hadronic particle collision (e.g., proton-proton), it is the partons themselves that collide. At sufficiently high energy, this can result in the disintegration of the composite particle. Drell and Yan suggested that a hadronic collision could be described as a weighted partonic collision [3].

1.2.1. Z Boson + Jet Production

In the SM, high energy proton-proton (*pp*) collisions can give rise to final states in which a Z^0 boson is produced in association with a parton. At the LHC, these "Z+jet" events are predominantly produced by quark exchange processes; primarily $\mathrm{q\bar{q}} \rightarrow \mathrm{Z}^0\mathrm{g}$ but also $\mathrm{qg} \rightarrow \mathrm{Z}^0\mathrm{q}$, both shown in Fig. 1.3.

A schematic representation of Z+jet production is shown in Fig. 1.4. The initial state partons carry a fraction of the incident proton's momentum, x_i. They interact to form a boosted state of mass M moving with momentum fraction $x = x_1 - x_2$ and

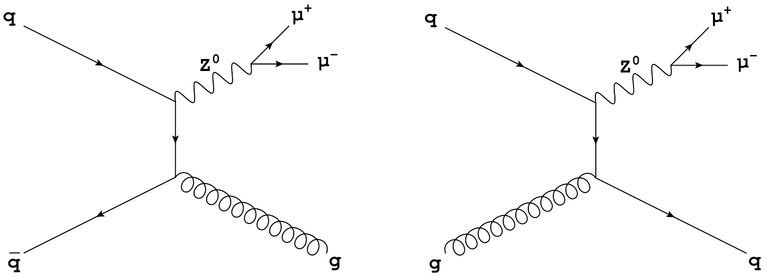


Figure 1.3: Two of the leading order t-channel Feynman diagrams for Z+jet production.

rapidity

$$y = \frac{1}{2} \ln \left[\left(E + p_z \right) / \left(E - p_z \right) \right], \tag{1.3}$$

where E is the energy and p_z is the component of the momentum along the beam axis. The state then “decays” into a two body final state with measured rapidities and transverse momenta (p_{T}).

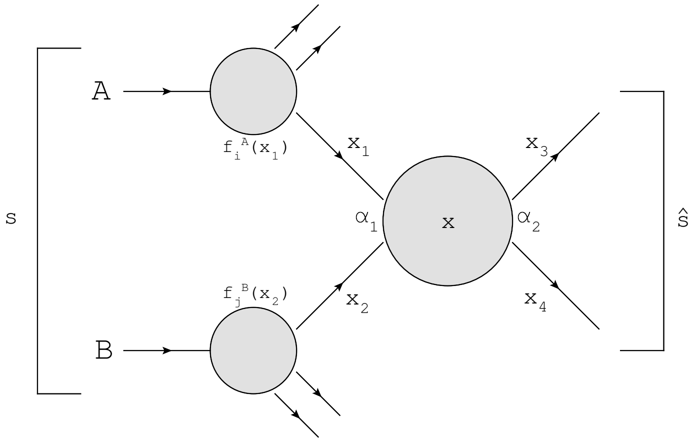


Figure 1.4: Schematic representation of the partons in Z+jet production, along with the kinematic variables that describe the scattering. The s represents the pp CM energy (7 TeV). All other variables are explained in the text. x_1 and x_2 represent the incoming partons, x_3 represents the Z^0 boson, and x_4 represents the outgoing jet.

The probability for a process to occur during a collision is proportional to the cross section σ, which is measured in units of barns ($1 \text{ b} = 10^{-28} \text{ m}^2$). The invariant cross section for inclusive hadronic reactions can be decomposed as [4],

$$\frac{E_3 E_4 \mathrm{d}^6\sigma}{\mathrm{d}^3\vec{p}_3 \mathrm{d}^3\vec{p}_4} \rightarrow \frac{\mathrm{d}^3\sigma}{\mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}\hat{t}} \sim \frac{1}{\hat{s}} \sum_{i,j} f_i(x_1) f_j(x_2) \frac{\mathrm{d}\hat{\sigma}_{ij}}{\mathrm{d}\hat{t}}(\hat{s}, \hat{t}, \hat{u}), \tag{1.4}$$

where p is the particle momentum; f is the parton distribution function (PDF) and is the probability that a parton of type i $\left(i = \mathrm{u}, \bar{\mathrm{u}}, \mathrm{d}, \bar{\mathrm{d}}, \mathrm{g}, \ldots\right)$ has a momentum fraction x; and $\hat{s}$, $\hat{t}$ and $\hat{u}$ are the Mandelstam variables. The first step of Eq 1.4 is an integration over a particle three-momentum (involving a delta function to conserve momentum). The partonic cross section takes the form of $\hat{\sigma} \propto (\alpha_s \alpha) \left|\mathcal{M}\right|^2 / \hat{s}$. The relevant Z+jet scattering amplitudes, $\left|\mathcal{M}\right|^2$, are listed in Table 1.1.

Table 1.1: Partonic scattering amplitudes for Z+jet production. The hat notation has been removed for simplicity.

Process	$\propto \left|\mathcal{M}\right|^2$
$q + \bar{q} \rightarrow \mathrm{Z}^0 + g$	$\frac{8}{9}\left(t^2 + u^2 + 2sm_3^2\right)/tu$
$q + g \rightarrow \mathrm{Z}^0 + q$	$-\frac{1}{3}\left(s^2 + u^2 + 2tm_3^2\right)/su$

Note that the partonic cross section is solely a function of the Mandelstam variables. In the CM frame they can be written as,

$$\begin{aligned} s &= (P_1 + P_2)^2 = (2p_1)^2 \\ t &= (P_1 - P_3)^2 = m_3^2 - \frac{s}{2}(1 - \cos\theta) \\ u &= (P_1 - P_4)^2 = -\frac{s}{2}(1 + \cos\theta), \end{aligned} \tag{1.5}$$

where P_i and p_i are the four- and three-momentum vectors, respectively (using the numbering scheme in Fig. 1.4, with the CM notation removed for simplicity). At high momentum transfers, partons/jets are essentially massless ($m_1 = m_2 = m_4 = 0$ in the last step of Eqs. 1.5).

At leading order, one can argue that the invariant cross section is a function of the PDFs, $\hat{s}$ and $\cos\theta^*$ (the CM polar scattering angle). This can be written explicitly by parameterizing the phase space in Eq. 1.4 as

$$\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}\hat{t} \propto \mathrm{d}(p^*)^2\mathrm{d}y_B\mathrm{d}\cos\theta^*, \tag{1.6}$$

where y_B is the rapidity of the CM or boosted system (see Fig. 1.4) and p^* is the Z^0 or jet momentum in the CM frame. One can then solve for the angular distribution of the Z^0, such that

$$\begin{aligned} \frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*} &\sim \int \sum_{i,j} \frac{f_i(x_1)}{x_1}\frac{f_j(x_2)}{x_2}\frac{\mathrm{d}\hat{\sigma}_{ij}}{\mathrm{d}\cos\theta^*}\mathrm{d}\left(p^*\right)^2\mathrm{d}y_B \\ x_{i,j} &= \frac{4\left(p^*\right)^2}{S}e^{\pm y_B}, \end{aligned} \tag{1.7}$$

Note that the variables p^* and y_B will be integrated over, which will play a major role in the accessible phase space (see Chapter 7).

1.3. Angular Distribution and Motivation

As previously mentioned, the relative Z^0 angular distribution is significant because it can be decoupled from the PDFs. A straightforward comparison between the measurement and pQCD predictions will highlight deviations from the theory, where anomalies may signify weaknesses in the SM at the new LHC energy regime. The SM posits that the angular distribution of the Z^0 is sensitive to the spin of the exchanged particle. To first-order, a spin-$\frac{1}{2}$ propagator produces an angular distribution of the form $\left(1 - |\cos\theta^*|\right)^{-1}$. However, the Z^0's angular distribution is predicted in the CM frame; the final state particle kinematics are measured in the lab frame. We therefore apply a Lorentz boost on the Z^0 and jet four-momentum vectors in order to transform them to the CM frame (discussed in Chapter 6). Note that, while the Z^0 may decay leptonically, we study only the muon decay channel because muons provide a very clean experimental signature [5].

The angular distribution of photon+jet, W+jet, and dijet events have been previously measured at the Tevatron CDF [6, 7] (see Fig. 1.5) and DØ [8] detectors, as well as the LHC CMS [9] and ATLAS [10] detectors. Although this is the first measurement of the Z+jet angular distribution, we expect similar results. Furthermore, the analysis in this dissertation includes regions of phase space that were not available to previous studies; we are able to probe larger values of $\cos\theta^*$ as a result of an increased reach in values of s and y. The data used in the analysis correspond to $\mathcal{L}_{int} \approx 36\ \text{pb}^{-1}$, which was taken in 2010 using the CMS detector.

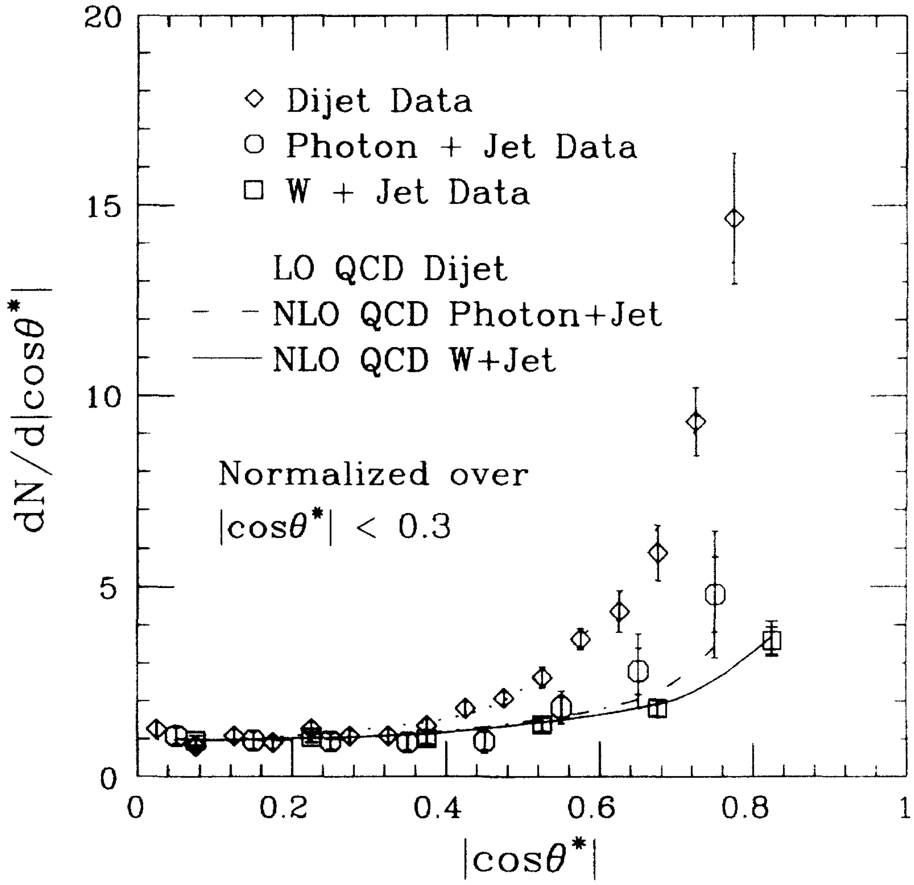


Figure 1.5: Selected $\mathrm{d}N/\mathrm{d}\,|\cos\theta^*|$ for CDF W+jet data (squares), compared to previously published measurements of $|\cos\theta^*|$ for dijet and photon+jet data. Next-to-leading-order QCD predictions are compared with the W+jet (solid curve) and the photon+jet (dashed curve) data. A leading-order QCD prediction (dotted curve) is compared to the dijet data. The data and theoretical predictions are all normalized to have an average value of 1 in the region $|\cos\theta^*| < 0.3$ [11].

CHAPTER 2

LHC AND THE CMS DETECTOR

2.1. LHC Accelerator

An in-depth description of the LHC can be found elsewhere [12]; however, details of the accelerator relevant to the analysis are summarized here.

The LHC is a 27 km long accelerator built to search for the origin of spontaneous symmetry breaking (i.e., the Higgs mechanism) and for physics beyond the SM with CM collision energies of up to 14 TeV. The number of events per second generated in LHC collisions is given by:

$$N_{event} = \mathcal{L}\sigma_{event}, \tag{2.1}$$

where σ_{event} is the cross section for the event under study and $\mathcal{L}$ the machine luminosity. The machine luminosity depends only on the beam parameters (see Table 2.1) and can be written for a Gaussian beam distribution as:

Table 2.1: LHC machine parameters, values vary in different operation scenarios.

Quantity	Variable	Design Value	2010 Value
Lorentz factor	γ	7461	3730
Revolution frequency	f	40 MHz	40 MHz
Number of bunches	k_B	2808	368
Protons per bunch	N_p	1.15×10^{11}	1.15×10^{11}
Transverse emittance	ϵ_n	3.75 μm rad	1.6 μm rad
Betatron function (IP)	β^*	0.55 m	3.5 m
Reduction factor	F	0.836	0.92
Luminosity	$\mathcal{L}$	10^{34} cm^{-2} s^{-1}	2×10^{32} cm^{-2} s^{-1}
Proton energy	E	7 TeV	3.5 TeV
Bunch separation		25 ns	75 ns
Number of collisions/crossing	n_c	≈ 20	≈ 5

$$\mathcal{L} = \frac{\gamma f k_B N_p^2}{4\pi\epsilon_n\beta^*}F, \tag{2.2}$$

where γ is the Lorentz factor, f is the revolution frequency, k_B is the number of bunches, N_p is the number of protons per bunch, ϵ_n is the normalized transverse emittance, β^* is the betatron function at the interaction point (IP), and F is the reduction factor resulting from the crossing angle. During the 2010 data collection period (that ran from May to November), the energy of each proton beam was 3.5 TeV, with the intention of upgrading to 5 TeV in the future. The design luminosity of $\mathcal{L} = 10^{34}\ \mathrm{cm^{-2}\,s^{-1}}$ results in approximately 1 billion pp interactions per second, whereas the peak luminosity for the 2010 run was $\mathcal{L} = 2\times10^{32}\ \mathrm{cm^{-2}\,s^{-1}}$ (see Fig. 2.1).

2.2. CMS Detector

An in-depth description of CMS can be found elsewhere [5]; however, details of the detector relevant to the analysis are summarized here.

The CMS collaboration uses a right-handed coordinate system, with the origin at the nominal IP, the x-axis pointing to the center of the LHC, the y-axis pointing up (perpendicular to the LHC plane), and the z-axis along the counterclockwise-beam direction. The polar angle, θ, is measured from the positive z-axis and the azimuthal angle, ϕ, is measured in the x-y plane. Also used as a coordinate is pseudorapidity, $\eta = -\ln[\tan(\theta/2)]$. Pseudorapidity is the ultra-relativistic (or massless) approximation of rapidity; it depends only on the polar angle and is preferred over θ since particle production is approximately constant as a function of rapidity.

The central feature of CMS is a superconducting solenoid. The silicon pixel and strip tracker, the crystal electromagnetic calorimeter (ECAL) and the brass/scintillator hadron calorimeter (HCAL) are within the field volume. Muons are measured in gas-ionization detectors embedded in the steel return yoke. In addition to the barrel and

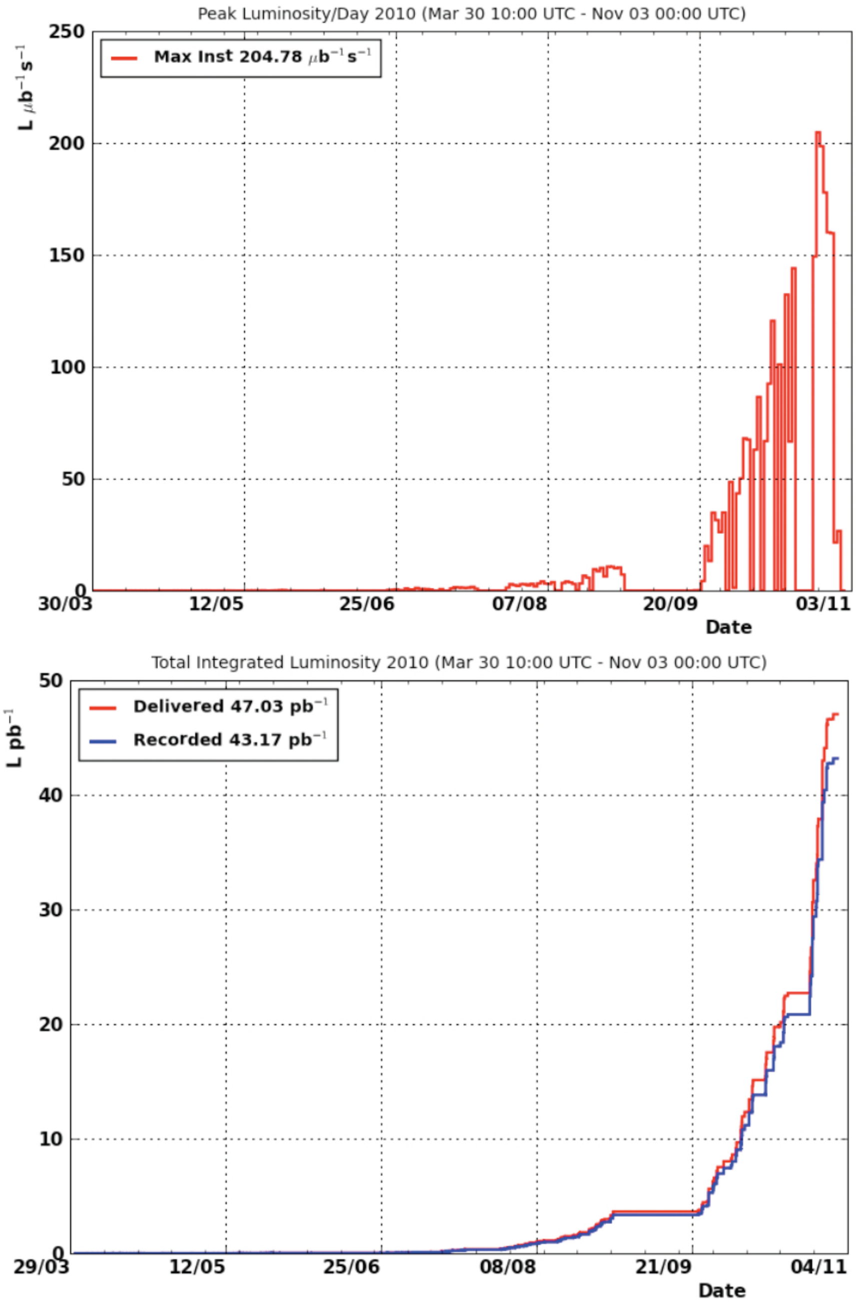


Figure 2.1: Instantaneous (top) and integrated (bottom) luminosity of the LHC during the 2010 run.

endcap detectors, CMS has extensive forward calorimetry. Diagrams of the CMS experiment are shown in Figs. 2.2 and 2.3. The overall dimensions of the CMS detector are a length of 21.6 m, a diameter of 14.6 m and a total weight of 12 500 tons.

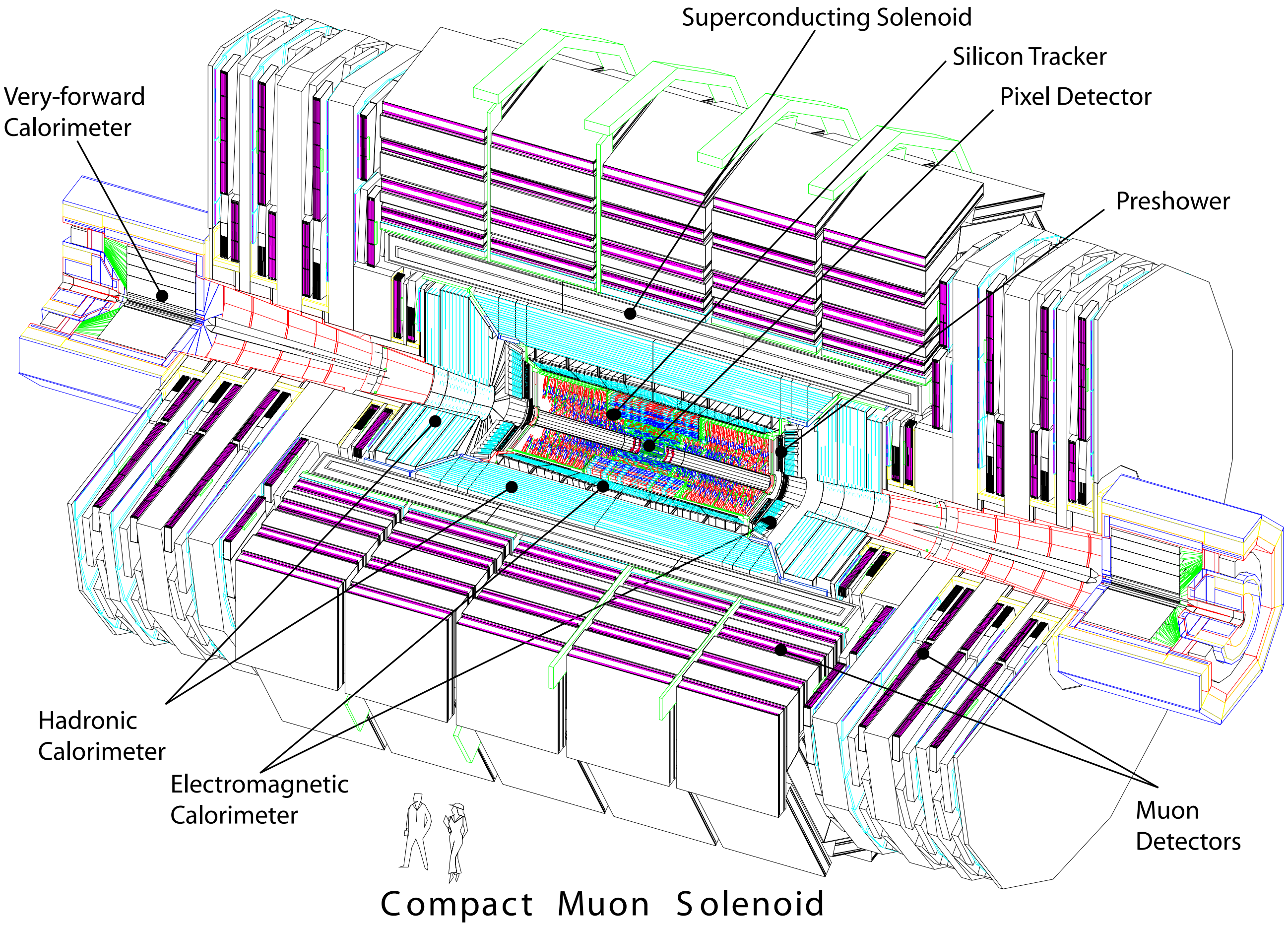


Figure 2.2: A drawing of the CMS experiment [5].

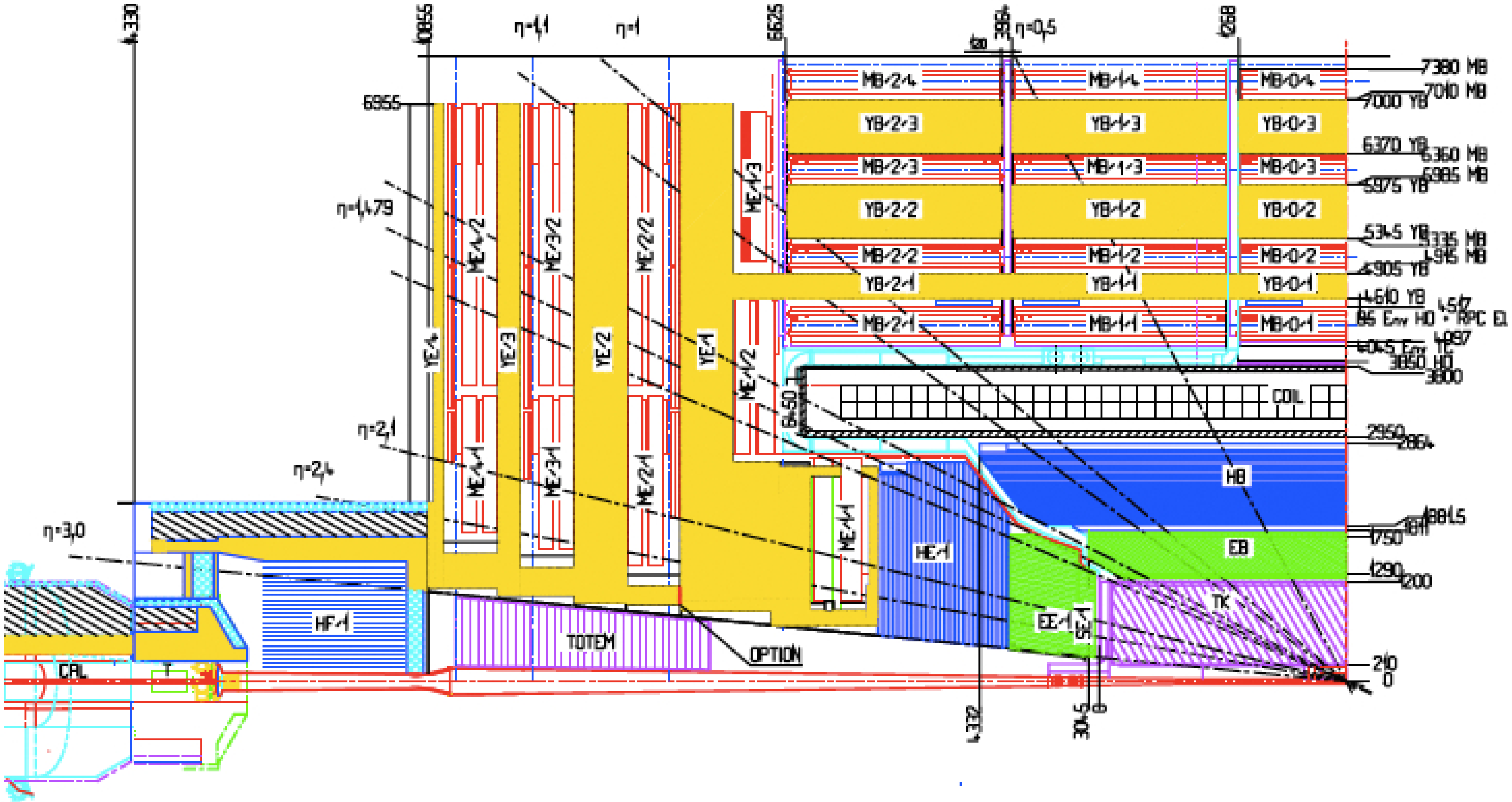


Figure 2.3: One-quarter longitudinal view of the CMS experiment. Dimensions are in units of mm and lines of constant pseudorapidity are superimposed [5].

2.2.1. Magnet

The CMS magnet is a 13 m long, 5.9 m inner diameter, 3.8 T superconducting solenoid. A high magnetic field was chosen to achieve good momentum resolution within a compact spectrometer, without making stringent demands on muon-chamber resolution and alignment. The bending power of the magnet is determined by the requisite performance of the muon system and by the need to unambiguously measure the sign for muons with a momentum of ≈ 1 TeV, which requires a momentum resolution of $\Delta p/p \approx 10\%$ at $p = 1$ TeV muons.

2.2.2. Tracking

The inner tracker is located in the magnetic field of the superconducting solenoid and measures charged particles within the range $|\eta| < 2.5$ (see Fig. 2.4). By considering the charged particle flux at various radii, three tracker regions can be identified:

- Closest to the IP, where the particle flux is the highest ($\approx 10^7$/s at $r \approx 10$ cm), CMS uses pixel detectors. The size of a pixel is $\approx 100 \times 150\ \mu\text{m}^2$, giving an occupancy of about 10^{-4} per pixel per bunch crossing.

- In the intermediate region ($20 < r < 55$ cm), the particle flux is low enough to enable use of silicon microstrip detectors with a minimum cell size of $10\ \text{cm} \times 80\ \mu\text{m}$, leading to an occupancy of $\approx 2-3\%$ per bunch crossing.

- In the outermost region ($r > 55$ cm), the particle flux has dropped sufficiently to allow the use of larger-pitch silicon microstrips with a maximum cell size of $25\ \text{cm} \times 180\ \mu\text{m}$, while keeping the occupancy to $\approx 1\%$.

The performance of the tracker (illustrated in Fig. 2.5) provides an impact parameter resolution of $\sim 15\ \mu$m and a p_{T} resolution of about 1.5% for 100 GeV particles. Tracking is a vital component to the reconstruction of collision products; therefore, we expand upon the track reconstruction algorithms used in CMS.

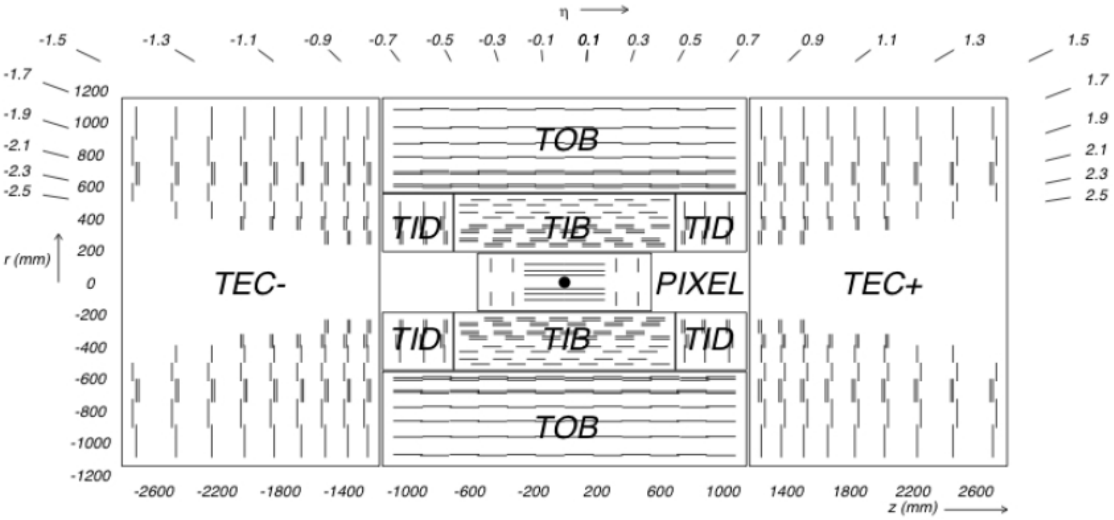


Figure 2.4: Schematic of a slice of the tracker parallel to the beam line. Lines of constant pseudorapidity are superimposed. The placement of the Tracker Inner Barrel (TIB), Tracker Inner Disks (TID), Tracker Outer Barrel (TOB) and Tracker End Cap (TEC) are also shown.

Track Reconstruction

Track reconstruction in a dense environment requires an efficient search for signals left by charged particles ("hits"). In the CMS tracker, this is simplified by the arrangement of sensitive modules in layers that are practically hermetic for a particle originating from the center of the detector. Another advantage is the fact that the magnetic field is almost constant in a large part of the tracker volume; a helical track model is adequate. For reconstruction purposes, a detailed distribution of passive material is replaced by an attribution of material to layers and simplifies the estimation of energy loss and multiple scattering.

The track reconstruction is decomposed into five logical parts [5]: hit reconstruction, which consists of the clustering of strips or pixels to estimate a position; seed generation; pattern recognition or trajectory building; ambiguity resolution; and final track fit.

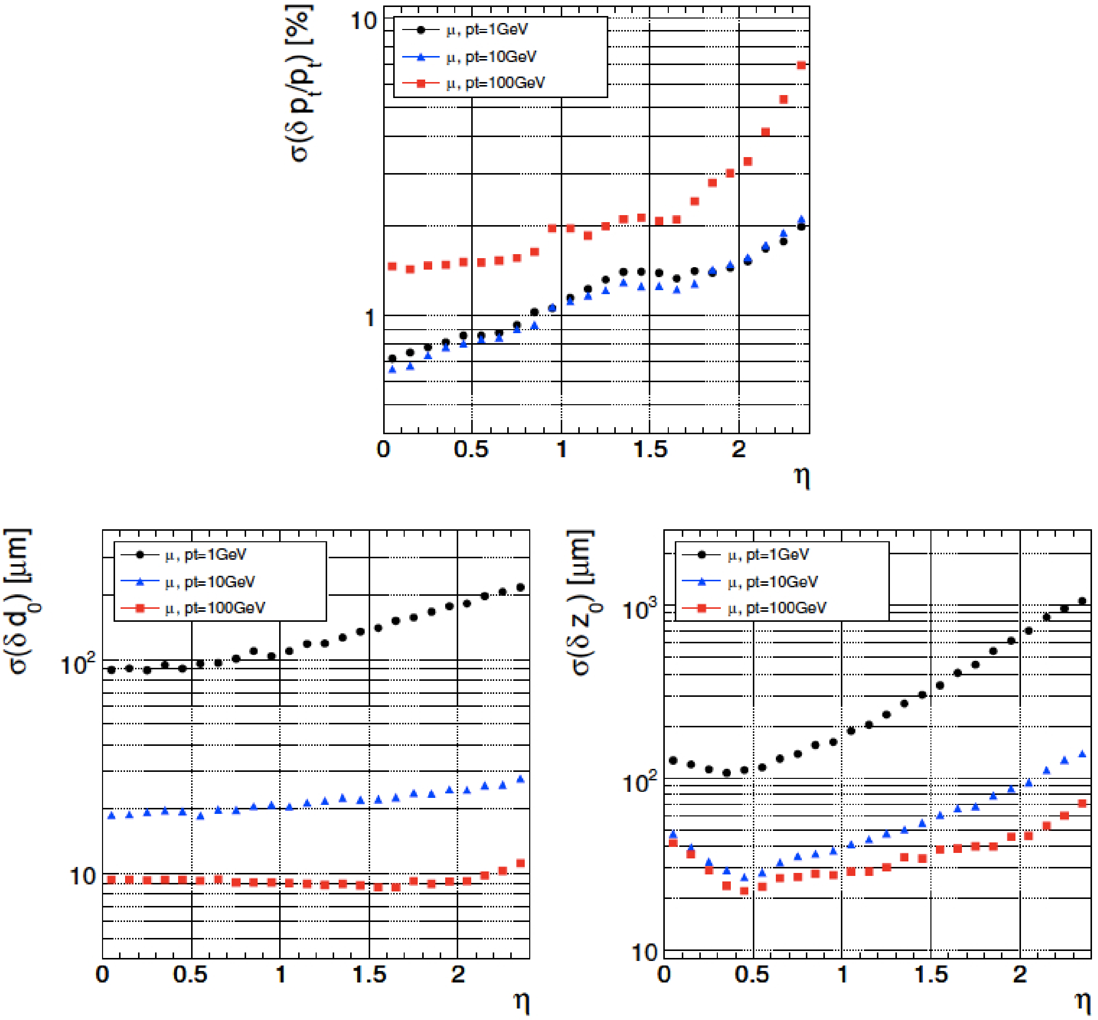


Figure 2.5: Resolution of several track parameters for single muons with p_T of 1, 10 and 100 GeV: (upper) transverse momentum, (lower-left) transverse impact parameter, and (lower-right) longitudinal impact parameter [5].

Strip Clustering

Clusters are reconstructed in the strip tracker by searching for a seed strip with a signal to noise ratio $S/N > 3$. Nearby strips are included in the cluster if they satisfy $S/N > 2$. The total signal size of the cluster must exceed $5 \times \sqrt{\sum \sigma_{\mathrm{rms}}^2}$, where σ_{rms} is the constituent strip's RMS-noise. The cluster position is usually determined from the centroid of the signal heights.

The cluster reconstruction algorithm for the pixel detector starts from a cluster seed, which defined as a pixel with $S/N > 6$. It then adds pixels adjacent to the cluster if they have $S/N > 5$, continuing this process until no more adjacent pixels are found (diagonally adjacent pixels are considered adjacent). Finally, the cluster is retained if its total charge has $S/N > 10.1$. The position of pixel clusters is estimated independently in both dimensions; it is derived from the relative charges of the pixels at the edges of the cluster and the associated reconstructed track angle.

The reconstruction inefficiency is defined as the fraction of simulated hits that do not have any associated reconstructed hit. For the pixel detectors this is below 0.5%. The fraction of reconstructed hits that is not associated with any simulated hit is less than 0.01%.

Seed Generator

Seed generation provides initial trajectory candidates for the full track reconstruction. In general, seeds are defined by pairs of pixel hits; the pixel detector is used because of its low occupancy. Seed creation is much more computationally intensive than just the two (minimum) hits used for its definition; it involves computation of the trajectory and construction of an error matrix. The time to generate a seed is about 0.3 ms/seed.

Pattern Recognition

The pattern recognition is derived from a combinatorial Kalman filter method. The filter proceeds iteratively from the seed, starting from a coarse estimate of the track parameters. A dedicated navigation component then determines which layers are compatible with the initial seed trajectory. Then, the trajectory is extrapolated to these layers according to the equations of motion of a charged particle in a magnetic field, accounting for multiple scattering and energy loss in the traversed material. Since several hits on the new layer may be compatible with the predicted trajectory, a trajectory candidate is created per hit. In addition, an extra trajectory candidate is created in which no measured hit is used; this accounts for the possibility that the track did not leave any hit on that particular layer. This fake hit is called an invalid hit.

Each trajectory is then updated with the corresponding hit according to the Kalman filter formalism. An update can be seen as a weighted mean of the predicted trajectory state and the hit; the weights attributed to the measured and predicted trajectories depend on their respective uncertainties. All resulting trajectory candidates are then grown in turn to the next compatible layer, and the procedure is repeated until either the outermost layer of the tracker is reached or a stopping condition is satisfied. To avoid an exponential increase in the number of combinations, a limited number of candidates are retained at each step by consider their normalized χ^2 and the number of valid and invalid hits.

Ambiguity Resolution

Ambiguities in track finding arise because a given track may be reconstructed starting from different seeds, or because a given seed may result in more than one trajectory candidate. These ambiguities must be resolved in order to avoid double counting of tracks.

The ambiguity resolution depends on the fraction of hits that are shared between two trajectories. For any pair of track candidates, this fraction is defined in the following way:

$$f_{\text{shared}} = \frac{N_{\text{shared}}^{\text{hits}}}{\min\left(N_1^{\text{hits}}, N_2^{\text{hits}}\right)}, \tag{2.3}$$

where N^{hits} is the number of hits in the track candidate. If this fraction exceeds a value of 0.5, the track with the least number of hits is discarded; if both tracks have the same number of hits, the track with the highest χ^2 value is discarded. The ambiguity resolution is applied twice: the first time on all track candidates resulting from a single seed, and the second time on the complete set of track candidates from all seeds.

Track Fitting and Smoothing

For each trajectory, the building stage results in a collection of hits and in an estimate of the track parameters. However, the full information is only available at the last hit of the trajectory and the estimate can be biased by constraints applied during the seeding stage. Therefore the trajectory is refitted using a least-squares approach, implemented as a combination of a standard Kalman filter/smoother. This filtering and smoothing procedure yields optimal estimates of the parameters at the surface associated with each hit and, specifically, at the first and the last hit of the trajectory. Estimates on other surfaces, e.g., at the impact point, are then derived by extrapolation from the closest hit.

2.2.3. Calorimetry

ECAL

The ECAL is a hermetic, homogeneous calorimeter consisting of 61 200 lead tungstate ($PbWO_4$) crystals mounted in the central barrel and sealed by 7324 crystals in each of the two endcaps (see Figs.2.6 and 2.7). The CMS collaboration has chosen $PbWO_4$ scintillating crystals because they have short radiation and Moliere lengths ($X_0 =$

0.89 cm and 2.2 cm, respectively), are fast emitters of light (80% of the light is emitted within 25 ns), and are radiation hard (up to 10 Mrad).

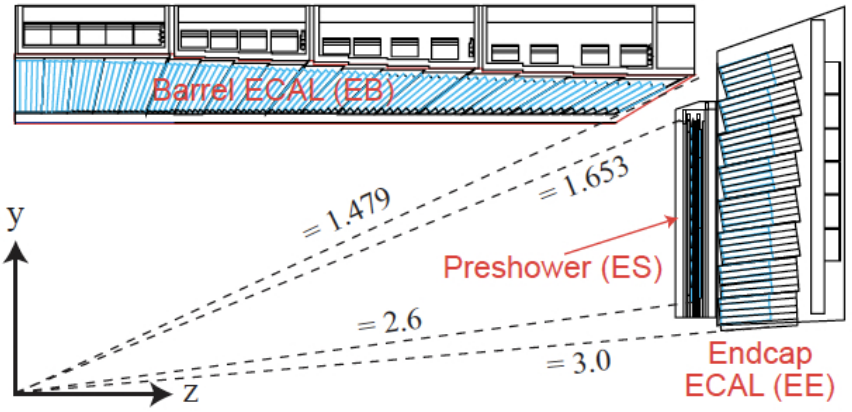


Figure 2.6: Transverse section through the ECAL, showing geometrical configuration. Lines of constant pseudorapidity are superimposed [5].

The barrel section has an inner radius of 129 cm. It is structured as 36 identical supermodules, each covering half the barrel length and corresponding to a pseudorapidity interval of $|\eta| < 1.479$. The crystals are quasi-projective (the axes are tilted at 3° with respect to the nominal vertex position) and cover 0.0174 (i.e., 1°) in $\Delta\eta$ and $\Delta\phi$.

Each endcap, at a distance of 314 cm from the vertex and covering a pseudorapidity range of $1.479 < |\eta| < 3.0$, is structured as two "Dees" consisting of semi-circular aluminum plates with units of 5×5 crystals (called supercrystals). The endcap crystals are also quasi-projective, but are arranged in an x-y grid. A preshower device is placed in front of the calorimeter over much of the endcap pseudorapidity range. The active elements of this device are two planes of silicon strip detectors, with a pitch of 1.9 mm, which lie behind disks of lead absorber at depths of 2 X_0 and 3 X_0.

The performance of a supermodule was measured in a test beam. Representative results on the energy resolution as a function of beam energy are shown in Fig. 2.8. The energy resolution (σ_E), measured by fitting a Gaussian function to the recon-

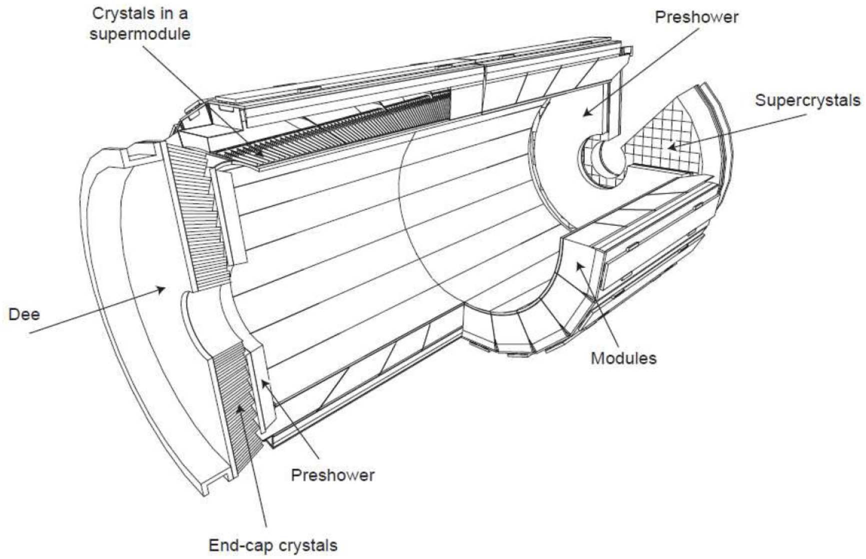


Figure 2.7: Schematic of the ECAL, showing the placement of crystals in the barrel and endcap. Also shown is the preshower [5].

structed energy distributions, has been parameterized as a function of energy,

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \tag{2.4}$$

where S is the stochastic term, N is the noise term and C is the constant term. The ECAL has an energy resolution of better than 0.5% for unconverted photons with transverse energies (E_{T}) above 100 GeV.

HCAL

The design of the HCAL is strongly influenced by the choice of magnet parameters; most of the calorimetry is located inside the magnet coil and surrounds the ECAL system (see Fig. 2.9). An important requirement of HCAL is to minimize the non-Gaussian tails in the energy resolution and to provide good hermeticity for the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) measurement. Hence, the HCAL design maximizes material inside the magnet coil in terms of interaction lengths. Brass has been chosen as the absorber material since it has a reasonably short interaction length, is

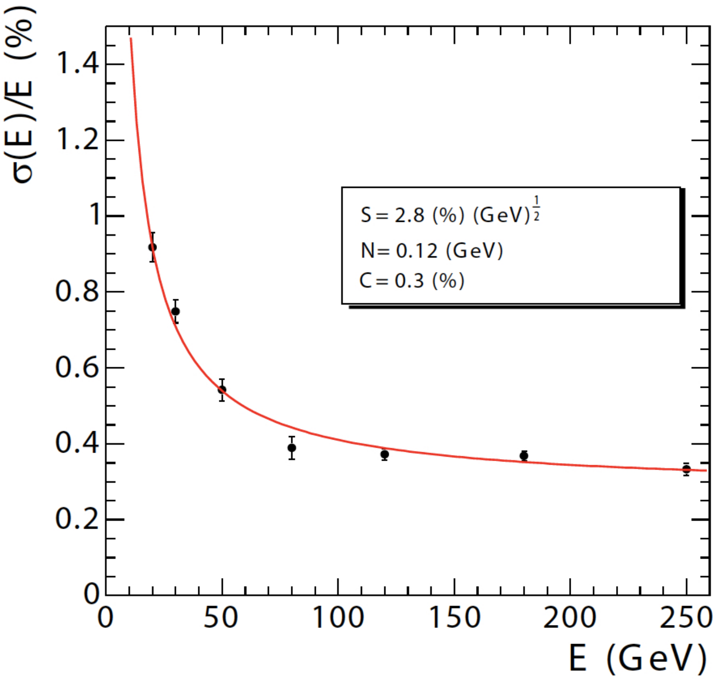


Figure 2.8: ECAL supermodule energy resolution, σ_E/E, as a function of electron energy as measured from a beam test. The energy was measured in an array of 3×3 crystals with electrons impacting the central crystal [5].

easy to machine, and is non-magnetic. Maximizing the amount of absorber before the magnet requires minimizing the amount of space devoted to the active medium. The tile/fiber technology makes for an ideal choice; it consists of plastic scintillator tiles read out with embedded wavelength-shifting fibers. These fibers are spliced to high-attenuation-length clear fibers outside the scintillator that carry the light to the readout system.

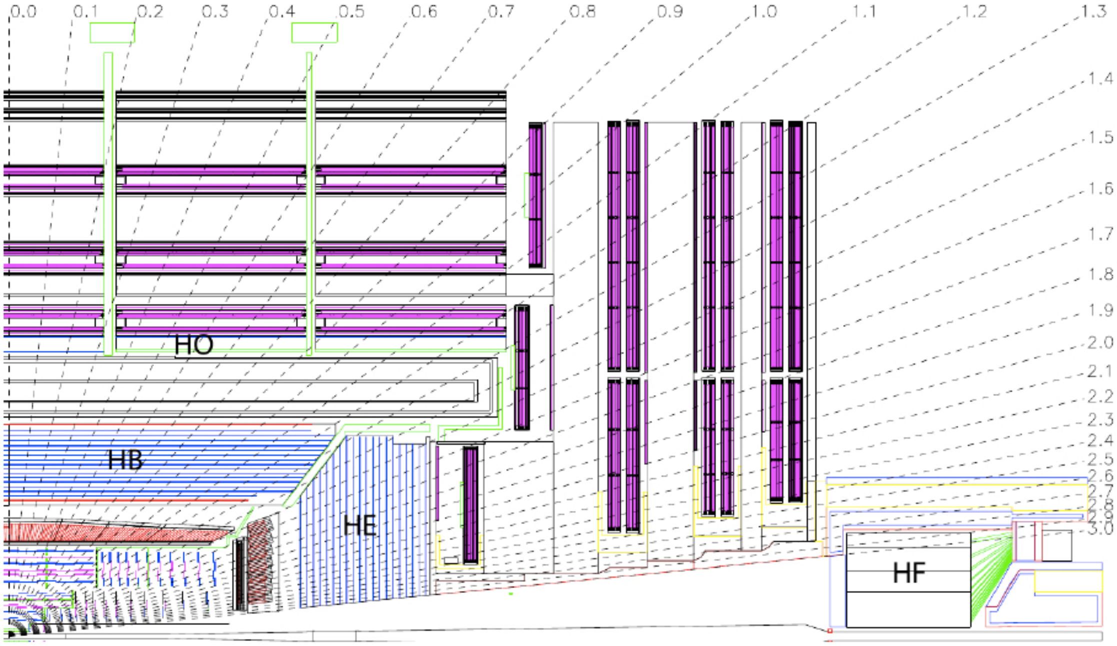


Figure 2.9: Longitudinal view of the CMS detector highlighting the locations of the HCAL barrel (HB), endcap (HE), outer (HO) and forward (HF). Lines of constant pseudorapidity are superimposed [5].

The hadron barrel consists of 32 towers per ($\Delta\phi = 5°$) wedge in the pseudorapidity region $|\eta| < 1.4$, resulting in 2304 towers with a $\Delta\eta \times \Delta\phi$ segmentation of 0.087×0.087. Each hadron endcap consists of fourteen η towers with 5° $\Delta\phi$ segmentation, covering the pseudorapidity region $1.3 < |\eta| < 3.0$. For the five outermost towers (at smaller η), the $\Delta\phi$ segmentation is 5° and the $\Delta\eta$ segmentation is 0.087. For the eight innermost towers, the $\Delta\phi$ segmentation is 10°, while the $\Delta\eta$ segmentation varies from 0.09 to 0.35. There are a total number of 2304 endcap towers.

For $|\eta| < 1.48$, the HCAL cells map on to 5×5 ECAL crystal arrays to form calorimeter towers projecting radially outwards from the IP. At larger values of η, the size of the towers increases and the matching ECAL arrays contain fewer crystals. Within each tower, the energy deposits in ECAL and HCAL cells are summed to define the calorimeter tower energies and subsequently used to provide the energies and directions of hadronic jets. The HCAL, when combined with the ECAL and tracker, measures jets with a resolution $\sigma_{p_\mathrm{T}}/p_\mathrm{T} \approx 10\%$ (see Fig. 2.10) using the particle-flow algorithm (discussed in Chapter 3).

2.2.4. Muon System

A schematic of the muon system is shown in Fig. 2.11. Three types of gaseous detectors are used to identify and measure muons. The choice of the detector technologies was driven by the large surface coverage required and by the different radiation environments. The drift tube (DT) chambers are used in the barrel region ($|\eta| < 1.2$), since the neutron induced background, the muon rate, and the residual magnetic field are low. In the two endcaps (where the parameters mentioned above are high), cathode strip chambers (CSCs) are deployed and cover the region up to $|\eta| < 2.4$. In addition to this, resistive plate chambers (RPCs) are used in both the barrel and the endcap regions. The RPCs provide a fast response with good time resolution but coarser position resolution; they can therefore identify the correct bunch crossing unambiguously. Details of the RPC system are covered elsewhere [5].

Drift Tubes

The barrel consist of 250 chambers organized into four stations at radii of approximately 4.0, 4.9, 5.9 and 7.0 m from the beam axis. Chambers in different stations are staggered so that a high-p_T muon produced near a sector boundary crosses at

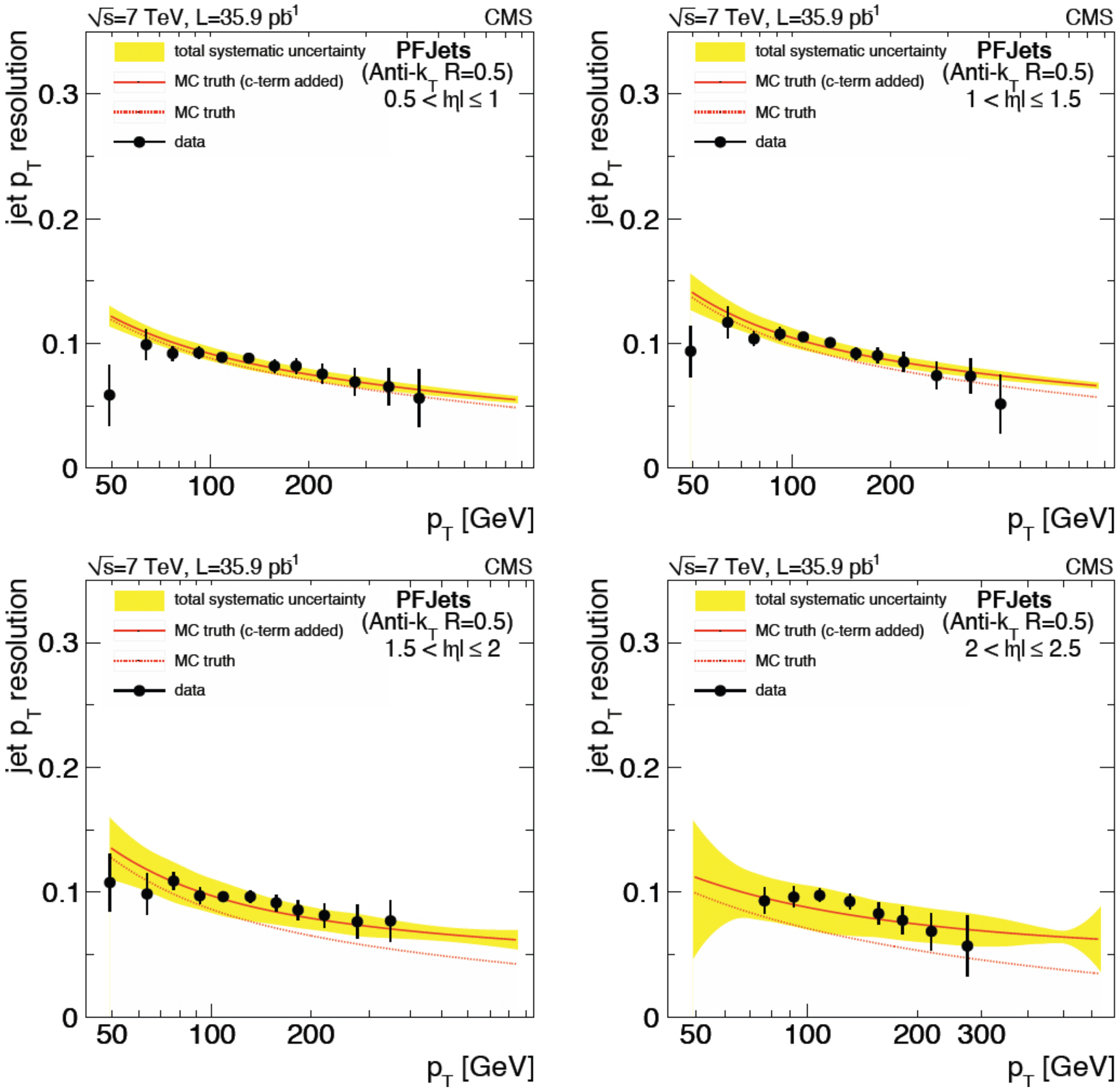


Figure 2.10: The jet p_{T} resolution as a function of the simulated jet p_{T} for jets in different η regions. The jets are reconstructed with the particle-flow algorithm (discussed in Chapter 3) [13].

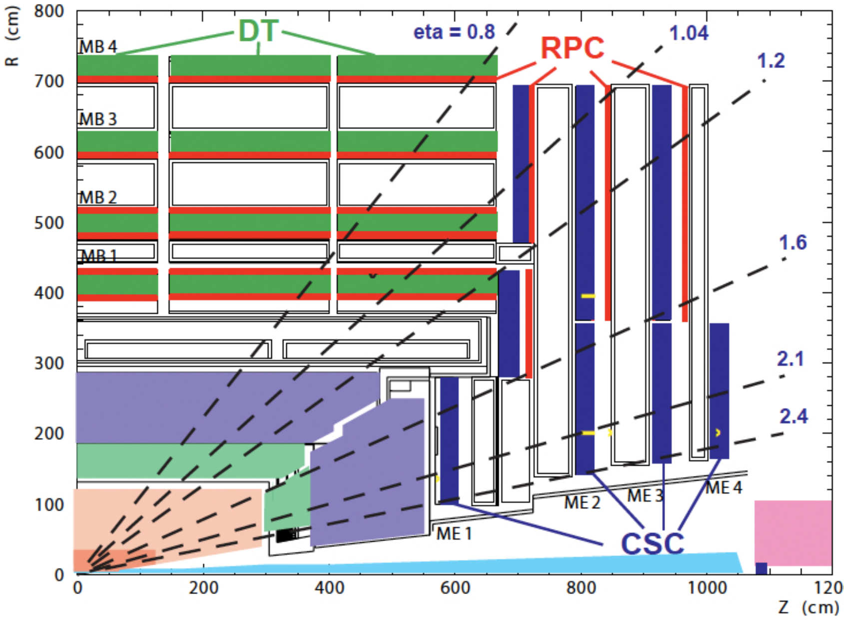


Figure 2.11: Layout of one quarter of the CMS muon system. The RPC system is limited to $|\eta| < 1.6$ in the endcap [5].

least three out of the four stations. In the innermost three station, chambers consist of twelve planes of aluminum drift tubes; the maximum drift length is 2.0 cm and the single point resolution is ≈ 200 μm. However, each station is designed to give a muon vector with a position precision better than 100 μm and an angular precision of approximately 1×10^{-3} rad. Depending on the station, a DT chamber can have one or two RPCs coupled to it. A high-p_{T} muon may cross up to six RPCs and four DT chambers, producing up to 44 measured points in the DT system from which a muon-track candidate can be built.

Cathode Strip Chambers

The endcap system consists of 468 CSCs divided into two endcaps, where most CSCs are overlapped in ϕ to avoid gaps in the muon acceptance. Each CSC is trapezoidal in shape and consists of six gas gaps, each gap having a plane of radial cathode strips and a plane of anode wires running almost perpendicularly to the strips. The gas ionization and subsequent electron avalanche caused by a charged particle traversing a chamber produces a charge on the anode wire and an image charge on a group of cathode strips. The signal on the wires, while fast, results in a coarse position resolution. A precise position measurement is made by determining the center-of-gravity of the charge distribution induced on the cathode strips. Each CSC measures up to six space coordinates (r, ϕ, z), where the spatial resolution provided by each chamber is typically about 200 μm. The angular resolution in ϕ is of the order 10×10^{-3} rad.

Muon Tracking

Centrally produced muons are measured three times: in the inner tracker, after the coil, and in the flux return yoke. The measurement of the momentum using only the muon system is determined by the muon bending angle at the exit of the magnet coil (taking the IP as the origin of the muon). The resolution of this measurement (labelled "muon system only" in Fig.2.12) is dominated by multiple scattering in the material located before the first muon station (for muons with $p_{\mathrm{T}} < 200$ GeV). The resolution of higher p_{T} muons is dominated by the chamber spatial resolution. For low-momentum muons, the best momentum resolution is given by the measurement obtained in the silicon tracker ("inner tracker only"). However, the muon trajectory beyond the return yoke can be extrapolated back to the beam-line and can be used to improve the muon momentum resolution at high p_{T} by combining the inner tracker and muon detector measurements ("full system"). Fitting muon system tracks to tracker tracks results in a p_{T} resolution between 1 and 5%, for p_{T} values up to 1 TeV.

2.3. Trigger and Data Acquisition

The total pp cross section at $\sqrt{s} = 14$ TeV is roughly 100 mb. At design luminosity, CMS will therefore observe a rate of approximately 10^9 inelastic collisions per second. The online event selection process ("trigger") must reduce $\approx 10^9$ interactions per

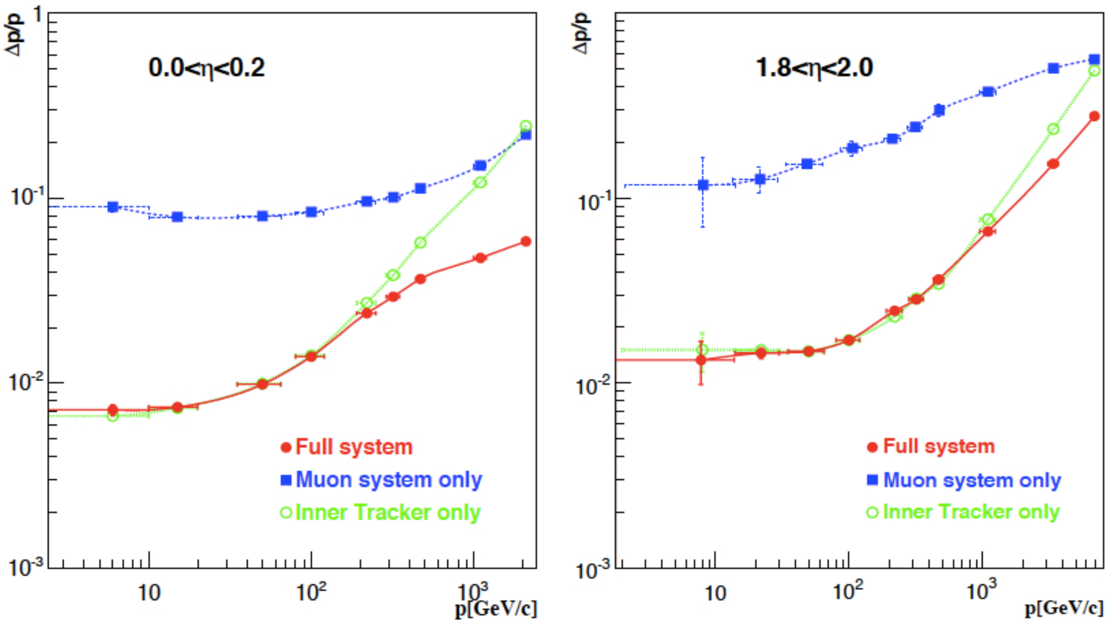


Figure 2.12: The muon momentum resolution versus p using the muon system only, the inner tracker only, or both (full system). Left shows the most central part of the barrel, $|\eta| < 0.2$ and right shows the endcap, $1.8 < \eta < 2.0$ [5].

second to no more than about 100 events per second for storage and subsequent analysis.

2.3.1. Level-1 Trigger

The size of the CMS cavern imposes a minimum transit time for signals from the detector electronics to reach the services cavern, where the Level-1 (L1) trigger logic is housed. The total time allocated for the transit is 3.2 μs. During this time, data from the detector are held in buffers while the L1 decides which events of interest to retain (only 1 in 1000 collisions are kept). Of the total latency, the time allocated to L1 trigger calculations is less than 1 μs.

The L1 triggers involve reduced granularity and resolution data from the calorimetry and muon systems. The decisions, formed by custom hardware processors, rely on the presence of trigger primitive objects (e.g., photons, electrons, muons, and jets) above various E_T or p_T thresholds. Decisions can also employ global sums of E_T and

$E_{\mathrm{T}}^{\mathrm{miss}}$. At design luminosity, the L1 rate will be limited to 100 kHz and is set by the average time to transfer full detector information through the readout system.

2.3.2. High Level Trigger

Upon receipt of a L1 trigger, detector data are transferred to readout buffers. After further signal processing, zero-suppression, and/or data-compression, the data are placed in memory for access by the data acquisition system. Each event has a size of about 1.5 MB. Upon event building, data from a given event are transferred to a processor. Each processor runs the same high-level trigger (HLT) software code to reduce the L1 output rate of 100 kHz to 100 Hz for mass storage.

Various strategies guide the development of the HLT code. Rather than reconstruct all possible objects in an event, only those objects and regions of the detector that are needed are reconstructed. Uninteresting events must be discarded as soon as possible, which leads to the idea of partial reconstruction and to the notion of virtual trigger levels (i.e., calorimeter and muon information is used first, followed by the use of the tracker data, and finally by the use of the full event information). Accepted events of interest are then ready to be reconstructed by the CMS software.

2.4. Computing

The CMS computing model needs to cover a broad range of activities including the design, construction, and calibration of the detector; the storage, reconstruction and analysis of data; and the support of a distributed computing infrastructure for physicists engaged in these tasks. The storage and processing power needed to analyze the data exceed the capabilities of most central computing systems. The CMS computing model is therefore highly distributed, with a basis of Grid technology, with a primary (Tier-0) center at CERN being supplemented by eight secondary (Tier-1) and over 50 tertiary (Tier-2) computing centers at national laboratories and universities worldwide. Tier-3 sites are relatively small computing installations that serve the needs of a local institution's users, and provide services and resources to CMS in a mostly opportunistic way. An in-depth description of the CMS computing model can be found elsewhere [14]; however, details relevant to the analysis are summarized here.

The CMS computing model is arranged in four tiers, which follows the flow of data shown in Fig. 2.13:

- A single Tier-0 center at CERN accepts data from the CMS data acquisition system, archives the data, and performs prompt reconstruction;
- The Tier-0 then distributes raw and processed data to Tier-1 centers in CMS collaborating countries. These centers provide services for data archiving, reconstruction, calibration, skimming and other data-intensive analysis tasks;
- A more numerous set of Tier-2 centers (smaller but with substantial CPU resources) provide capacity for analysis, calibration activities and simulation. Tier-2 centers rely upon Tier-1s for access to large datasets and secure storage of the new data they produce;
- Tier-3 centers provide interactive resources for local groups and additional computing capacity for the collaboration.

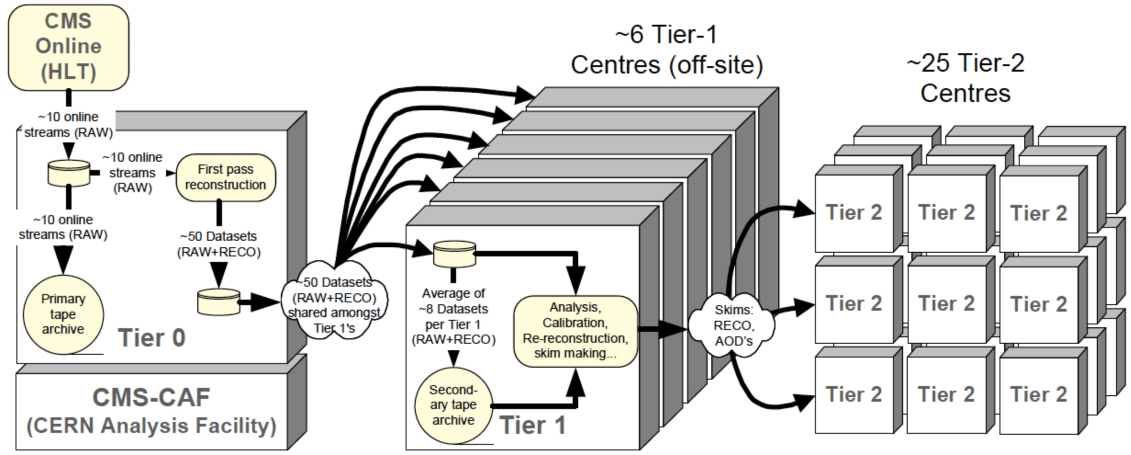


Figure 2.13: Schematic showing the flow of data from the CMS detector (during the design phase) through the various tiers of computing centers. The details of the figure are discussed in [14].

CHAPTER 3

EVENT RECONSTRUCTION

3.1. CMS Software

The overall collection of CMS software (CMSSW) is built around a framework that consists of simulation, calibration and alignment, and reconstruction modules. Each module processes event data so that physicists can perform analysis[1]. An in-depth description of CMSSW can be found elsewhere [5]; however, event reconstruction details relevant to the analysis are summarized here.

Reconstruction (RECO) is the operation of constructing physical quantities from the raw data collected from the experiment. The RECO process can be divided into three steps: local RECO within an individual subdetector module, global RECO within a whole subdetector, and the combination of these reconstructed objects to produce higher-level objects.

The units providing local RECO in a subdetector module can use either real data (i.e., from the data acquisition system) or simulated data. In either case, the output from the units are called "RecHits." Local RECO varies depending on the subdetector:

- In the tracker, local RECO algorithms search for strips/pixels with a signal exceeding a threshold, and use these as seeds for clusters. Clusters are built by adding neighboring strips/pixels.
- In the ECAL/HCAL, local RECO identifies the position, time of arrival, and energy of localized electromagnetic/hadronic energy depositions.
- In the muon DTs, local RECO provides the position of a muon hit in a drift cell, determined from the drift time measurement and the effective drift velocity.

---

[1] The analysis was performed using CMSSW 4_2_X

Three-dimensional track segments within a superlayer are built from hits in each component layer.

- In the muon CSCs, local RECO provides position and time of arrival of a muon hit from the distribution of charge induced on the cathode strips. Two-dimensional hits are obtained in each layer, and these can be combined to create three-dimensional track segments within each chamber.
- In the muon RPCs, local RECO gives the position of a muon hit from the position of clusters of strips.

In the global RECO step, information from the different modules of a subdetector are combined, although information from separate subdetectors is kept separate. For example, tracker RecHits are used to produce reconstructed charged particle tracks and muon system RecHits are used to produce candidate muon tracks.

The final RECO step combines reconstructed objects from individual subdetectors to produce higher-level objects suitable for the HLT or for physics analysis. For example, tracks in the tracker and tracks in the muon system are combined to provide final muon candidates, or matching ECAL and HCAL clusters are combined into jet candidates.

The analysis uses the CMS particle-flow algorithm; a modification of CMS high-level reconstruction.

3.2. Particle-Flow Algorithm

In contrast to the standard reconstruction, the particle-flow (PF) algorithm aims at reconstructing stable particles (i.e., electrons, muons, photons, charged and neutral hadrons) through a combination of all CMS subdetectors. Broadly speaking, reconstruction occurs using tracks and ECAL energy clusters for electrons, tracker and muon system tracks for muons, tracks and ECAL and HCAL clusters for charged hadrons, and ECAL/HCAL energy clusters for photons/neutral hadrons. A detailed description of the PF algorithm is discussed elsewhere [15, 16]. We briefly describe the algorithm, then discuss jet reconstruction in detail.

Description of the PF Algorithm

A given particle is expected to give rise to several PF elements in the various CMS subdetectors: one charged-particle track, and/or several calorimeter clusters, and/or one muon track. These elements must be connected to each other by a link algorithm to fully reconstruct a single particle (while avoiding double counting from different subdetectors). The algorithm then produces "blocks" of linked elements.

The PF algorithm starts by looking for muons. In CMS, muons are first reconstructed independently in the silicon tracker ("tracker track") and in the muon spectrometer ("standalone track"). A *global muon* reconstruction is then performed: starting from a standalone track in the muon system, a matching tracker track is found and a global muon track is fit by combining hits from both. A detailed explanation of muon reconstruction in CMS can be found elsewhere [17].

For each PF block, the algorithm proceeds as follows. First, each global muon gives rise to a "particle-flow muon" if its combined momentum is compatible with the momentum determined solely by the tracker (within three standard deviations). The corresponding track is removed from the block.

After finding muons, the algorithm then searches for electrons. Each track of the block is submitted to a pre-identification stage which exploits the tracker as a pre-shower: electrons tend to give rise to short tracks and to lose energy by Bremsstrahlung in the tracker layers on their way to the calorimeter. Pre-identified electron tracks are refit with a Gaussian-sum filter in an attempt to follow their trajectories all the way to the ECAL. A final identification is performed with a combination of a number of tracking and calorimetric variables. Each identified electron gives rise to a “particle-flow electron.” The corresponding track and ECAL clusters (including all ECAL clusters identified as Bremsstrahlung photons) are removed from the block.

Tighter quality criteria are then applied to the remaining tracks: it is required that the relative uncertainty on the measured p_{T} be smaller than the relative calorimetric energy resolution expected for charged hadrons. In hadronic jets, 0.2% of the tracks are rejected by this requirement. While about 90% of these are fake tracks, the energy of the remaining tracks originating from real particles is not lost as it is measured independently by the calorimeters. The remaining elements may give rise to charged hadrons, photons, neutral hadrons, and (more rarely) to additional muons.

A track can be directly connected to a number of ECAL and HCAL clusters. The detection of the neutral particles in the block (i.e., photons and neutral hadrons) involves a comparison between the momentum of the tracks and the energy detected in the calorimeters. Several tracks can be linked to the same HCAL cluster, in which case the sum of their momenta is compared to the calorimetric energy. However, if a track is linked to several HCAL clusters only the link to the closest cluster is kept for the comparison. A track can also be linked to more than one ECAL cluster, and the link to the closest cluster is again kept.

Any additional ECAL clusters might come from hadronic shower fluctuations, in which case the links ought to be preserved to avoid double counting of the hadron energy. Conversely, if these ECAL clusters arise from overlapping photons, the links ought to be ignored to allow for photon detection. Therefore, the ECAL clusters connected to any of the tracks under consideration are first ordered according to their distance to the closest track. The ordered list is then scanned and the corresponding link is kept as long as the total calorimetric energy remains smaller than the total charged-particle momentum.

These tracks gives rise to a “particle-flow charged hadron,” the momentum and energy of which are taken directly from the track momentum (under the charged pion mass hypothesis). If the calorimetric energy is compatible with the track momentum within measurement uncertainties, the charged-hadron momenta are redefined by a fit of the measurements in the tracker and the calorimeters. The combination is relevant at very high energies and/or large pseudorapidities, for which the track parameters are measured with degraded resolutions.

However, the energy of the closest ECAL and HCAL clusters may be significantly larger than the charged-particle momentum. If the relative energy excess is found to be larger than the expected calorimeter energy resolution, it gives rise to a “particle-flow photon”, and possibly to a “particle-flow neutral hadron”. If the excess is larger than the total ECAL energy, a photon is created with this ECAL energy and a neutral-hadron is created with the remaining part of the excess. Otherwise, the excess gives rise only to a photon. The precedence given in the ECAL to photons over neutral hadrons is justified by the observation that 25% of the jet energy is carried by photons, while neutral hadrons leave only 3% of the jet energy in the ECAL.

The remaining ECAL and HCAL clusters, either originally not linked to any track or for which the link was disabled, give rise to PF photons and PF neutral hadrons, respectively. It is these hadrons that will ultimately make up PF jets. Since the analysis pertains to the production of Z^0 and jets, a brief overview of the jet definition is instructive.

3.3. Jets

While a parton is well-defined at tree-level, from a computational and experimental point of view the concept is ambiguous. In pQCD, partons have divergent fragmentation probabilities that make the definition of a parton difficult [18]. Therefore, since jets are the representations of partons, one must introduce a set of rules for how to group particles into jets. A good jet definition should be applicable to partonic calculations, to parton showering, and to experimental measurements. Because of the variety of parton origins and dynamics, many jet definitions have been developed and used throughout the years. In the following section, we borrow heavily from Gavin Salam's review of jet finding at hadron colliders [18].

3.3.1. Jet Algorithms

Jet algorithms usually involve one or more parameters that indicate how close two particles must be for them to belong to the same jet. Additionally, they define a recombination scheme that indicates what momentum to assign to the combination of two particles (the simplest being the four-vector sum). The general properties of a jet definition was set out in the "Snowmass accord", which states that a jet definition should [18]

1. Be simple to implement in an experimental analysis;
2. Be simple to implement in theoretical calculations;
3. Be defined at any order of perturbation theory;
4. Yield finite cross sections at any order of perturbation theory;
5. Yield a cross section that is relatively insensitive to hadronization.

Jet algorithms can be divided into two broad categories. In the first category are those focused on cones; top-down algorithms that rely on the idea that QCD fragmentation/hadronization does not drastically affect an event's energy flow. In the second category are sequential recombination algorithms; bottom-up algorithms that repeatedly recombine the closest pair of particles, according to some distance measure motivated by the divergent structure of QCD matrix elements.

Cone Algorithms

Most modern cone algorithms are iterative cone (IC): a seed particle i sets some initial direction and one sums the momenta of all particles j within a cone of radius R around i, i.e., taking all j such that

$$\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 < R^2. \tag{3.1}$$

The direction of the resulting sum is then used as a new seed, and the procedure is iterated until the direction of the resulting cone is stable. Cone algorithms differ in their consideration of what to use as seeds and how to handle overlapping cones from distinct seeds (i.e., cones that share particles).

One approach is to find all possible stable cones by iterating over all particles and performing a split-merge procedure. The procedure merges a pair of cones if more than some fraction of the softer cone's energy is shared with the harder cone; otherwise the shared particles are assigned to the cone to which they are closer. Therefore, IC split-merge algorithms create all possible stable cones (protojets) then apply the following procedure [18],

1. Take the protojet with the largest p_{T} (i.e., the hardest protojet) and label it a.
2. Find the next hardest protojet that shares particles (i.e., overlaps) with a and label it b. If no such protojet exists, then remove a from the list of protojets (adding it to the list of final jets) and repeat from step 1.
3. Determine the total p_{T} of the particles shared between the two protojets.
    - If $p_{\mathrm{T}}^{shared}/p_{\mathrm{T}}^{b} > f$, where f is a free parameter known as the overlap threshold, replace protojets a and b with a single, merged protojet.

  - Otherwise split the protojets, assigning the shared particles to the protojet whose axis is closer.

4. Repeat from step 1, so long as there are protojets left.

Unfortunately, the outcome heavily depends on the initial set of stable cones. In fact, the addition of soft seed particles can lead to a different set of final jets. These algorithms are said to lack infrared and/or collinear safety.

Infrared and Collinear Safety

Infrared and collinear (IRC) safety means that, if one modifies an event by a collinear splitting or the addition of a soft emission, the set of final jets should remain unchanged. IRC safety is a desired property because [18]:

- A hard parton undergoes many collinear splittings as part of the fragmentation process. Additionally, there is always some emission of soft particles in QCD events. The final set of jets should not depend on these effectively random collinear splittings and soft emissions (point 5 of the Snowmass accord).

- In pQCD calculations, soft emissions and collinear splittings are associated with divergent tree-level matrix elements. There are also corresponding divergent loop matrix elements that enter with the opposite sign. Normally, the two sources of divergence cancel. However, for IRC unsafe jet algorithms, the tree-level splittings may lead to one set of jets, while the loop diagrams may lead to another, thus breaking the cancellation and leading to infinite cross sections (point 4 of the Snowmass accord).

Examples of IR unsafety are illustrated in Fig. 3.1. An event with two hard partons as seeds create two stable cones and two jets (a). The same occurs in the (negative) infinite loop diagram (b). In diagram (c), however, an extra soft gluon has been emitted; the gluon provides a new seed and creates a new stable cone that contains both hard partons (as long as they have similar momenta and are separated by less than $2R$). The stable cone overlaps with the two original ones and the result of the split-merge procedure is that only one jet is found. After pQCD integration over the virtual/real soft gluon momentum, the two-jet and one-jet cross sections each get non-canceling infinite contributions.

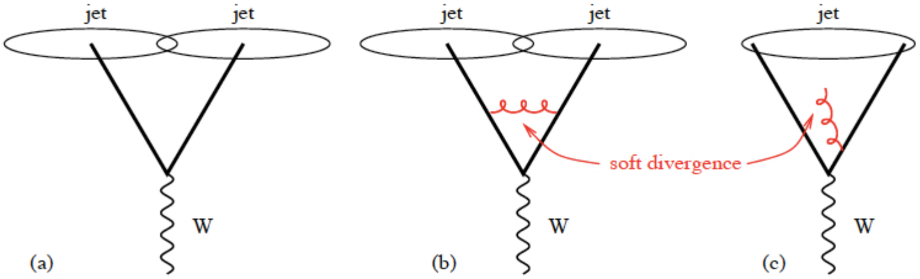


Figure 3.1: Configurations illustrating IR unsafety of IC split-merge algorithms. The addition of a soft gluon converts the event from having two jets (a and b) to just one jet (c). The explicit angular structure is shown [18].

One solution to the IRC safety issue avoids the use of seeds and iterations, and instead uses *all possible* stable cones. A modern implementation of this algorithm is SIScone (seedless infrared-safe cone algorithm) [18]. However, because it is not used in the analysis, it will not be discussed in this dissertation.

Sequential Recombination Algorithms

Sequential recombination algorithms go beyond just finding jets; they implicitly assign a clustering sequence to an event, which is closely connected with the probabilistic nature of parton branching. A simple example is the k_T algorithm (created for e^+e^- experiments) and is formulated as follows [18]:

1. For each pair of particles i, j calculate the distance

$$y_{ij} = \frac{2\min\left(E_i^2, E_j^2\right)\left(1 - \cos\theta_{ij}\right)}{Q^2}, \tag{3.2}$$

where Q is the total energy in the event, E is the energy of the particle, and θ_{ij} is the angle between particles i and j. In the collinear limit ($\theta_{ij} \ll 1$), the numerator just reduces to $[\min(E_i, E_j)\,\theta_{ij}]^2$ which is just E_i relative to j.

2. Find the minimum $y_{ij} \to y_{\text{min}}$.
3. If y_{min} is below some jet resolution threshold y_{thr}, then recombine i and j into a single new particle (or pseudojet) and repeat from step 1.
4. Otherwise, declare all remaining particles to be jets and terminate the iteration.

The k_{T} algorithm is infrared and collinear safe, because soft particles will get recombined right at the start of the clustering (as will collinear particles). The motivation for Eq. 3.2 is the pQCD splitting probability for one parton k to go into two, i and j, in the limit where either i or j is soft and both are collinear to each other [18],

$$\frac{\mathrm{d}P_{k\to ij}}{\mathrm{d}E_i \mathrm{d}\theta_{ij}} \sim \frac{\alpha_s}{\min(E_i, E_j)\,\theta_{ij}}. \tag{3.3}$$

Two issues arise in *pp* collisions. First, the total energy is not well defined (because non-hard-scatter particles may get lost down the beam pipe). Second, the divergences in the QCD branching probability are not just between pairs of outgoing particles, but also between an outgoing particle and the incoming beam direction. The hadronic generalization of the k_{T} algorithm is [18]

$$\begin{aligned} d_{ij} &= \min\big(p_{\mathrm{Ti}}^{2\rho}, p_{\mathrm{Tj}}^{2\rho}\big)\,\frac{\Delta R_{ij}^2}{R^2} \\ d_{iB} &= p_{\mathrm{Ti}}^{2\rho}, \end{aligned} \tag{3.4}$$

where d_{iB}, the distance between particle i and the beam, acts as a new threshold and R plays its analogous role from the cone algorithms. The ρ parameter can be set to 1 for the k_T algorithm; 0 for the Cambridge/Aachen algorithm, which is not discussed in this dissertation; and -1 for the the anti-k_T algorithm, which is used for the analysis. If $\rho = -1$, the algorithm favors clusterings that involve hard particles rather than soft particles (k_T algorithm), or energy-independent clusterings (Cambridge/Aachen). This ultimately means that the jets grow outwards around hard seeds. However, since the algorithm still involves a combination of energy and angle in its distance measure, it is a collinear-safe growth. The result is an IRC safe algorithm that gives circular hard jets. Figure 3.2 illustrates the jets that are produced with four common IRC-safe algorithms.

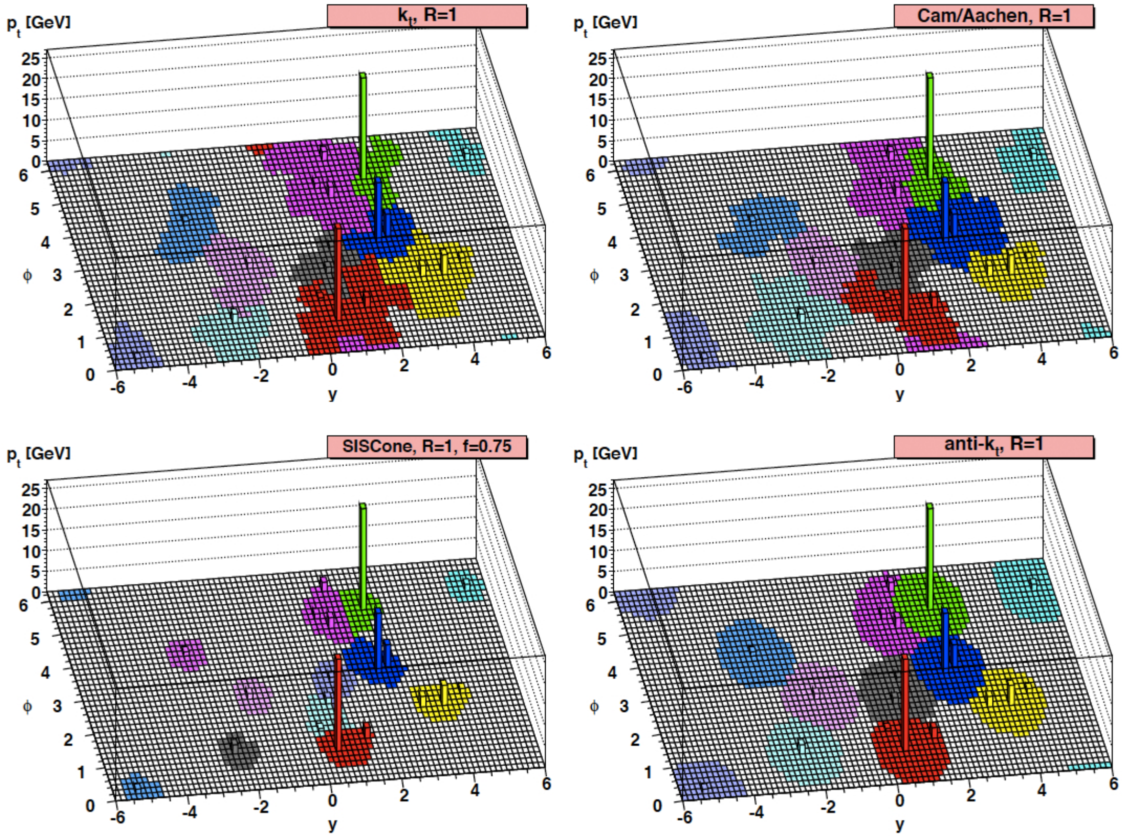


Figure 3.2: Hard partons and soft particles clustered with four different jet algorithms, illustrating the areas of the resulting hard jets [18].

For the analysis, hadronic PF particles are clustered into "particle-flow jets" by the anti-k_T algorithm [19] with a size parameter of $R = \sqrt{\eta^2 + \phi^2} = 0.5$. The raw jet energies are corrected to establish a relative uniform response of the calorimeter in η and a calibrated absolute response in p_T; this is known as the jet energy scale (JES) [13].

3.3.2. Jet Energy Scale

In general, the jet energy measured is different from the true particle-level jet energy. The difference is caused by the non-uniform and non-linear response of the calorimeters, with further complications arising from the presence of electronic noise and additional pp collisions in the same bunch crossing. To make measurements of jet energy as accurate possible, corrections are made to the energy scale of reconstructed jets.

Two primary corrections are made via comparisons to simulation (C_sim). The first is a relative correction that removes variations versus jet η relative to a uniform region of the detector. The second correction is an absolute correction which removes variations versus jet p_T. Other adjustments include an offset correction (C_offset) that removes the extra energy as a result of noise and pile-up and a residual correction (C_rel and C_abs) that accounts for the small differences between data and simulation. The various components are applied in sequence as described by the equation below:

$$p_\mathrm{T}^\mathrm{corrected} = p_\mathrm{T}^\mathrm{uncorrected} \times C_\mathrm{offset}(p_\mathrm{T}^\mathrm{raw}) \times C_\mathrm{sim}(p'_\mathrm{T}, \eta) \times C_\mathrm{rel}(\eta) \times C_\mathrm{abs}(p''_\mathrm{T}), \tag{3.5}$$

where p_T^{raw} is the uncorrected jet transverse momentum, p'_T is the momentum of the jet after applying the offset correction and p''_T is the momentum of the jet after all previous corrections. In the following sections, each component of the jet energy calibration will be discussed briefly. The total jet energy correction factors are shown in Fig. 3.6.

Offset Correction

The offset correction is the first step in the chain of factorized corrections. Its purpose is to subtract the energy not associated with the high-p_{T} scattering. An average p_{T} per unit area (ρ) in the event is estimated, which characterizes the soft jet activity and is a combination of the underlying event, the electronics noise, and the pile-up (see Fig. 3.3). The correction factor takes the form of

$$C_{\mathrm{area}}\left(p_{\mathrm{T}}^{\mathrm{raw}}, A_j, \rho\right) = 1 - \frac{\left(\rho - \langle \rho_{\mathrm{UE}} \rangle\right) A_j}{p_{\mathrm{T}}^{\mathrm{raw}}}, \tag{3.6}$$

where A_j is the jet area and $\langle \rho_{\mathrm{UE}} \rangle$ is the average p_{T}-density component attributed to the underlying event and electronics noise.

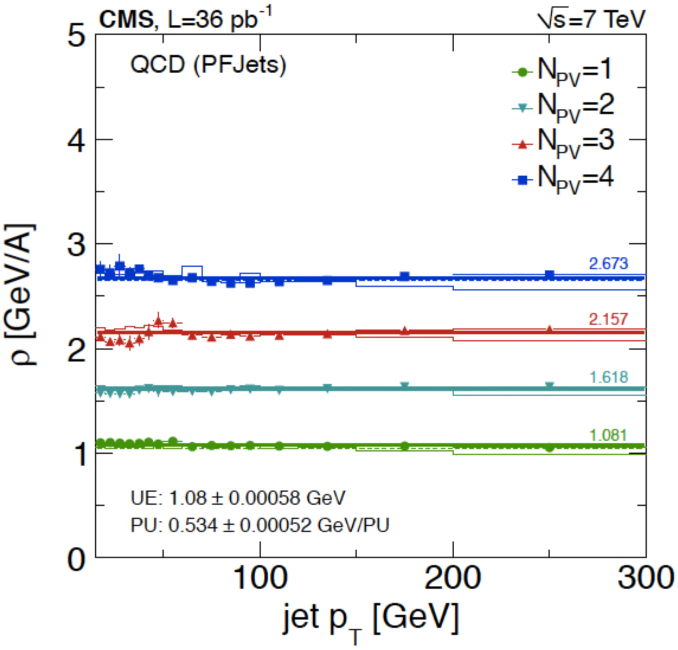


Figure 3.3: Pile-up and underlying event p_{T}-density as a function of the leading jet p_{T} in a QCD multijet simulation for various pile-up conditions (here N_{PV} denotes the number of reconstructed vertices, and A denotes the unit area in the y-ϕ space) [13].

Simulation Calibration

The simulation calibration is derived from simulation and corrects the energy of the reconstructed jets such that it is equal on average to the energy of the generated particle jets. Each reconstructed jet is spatially matched in η-ϕ space with a generated particle jet. Figure 3.4 shows the simulation jet energy correction factor for three jet types *vs.* η and p_{T}; CALO jets are created only from calorimeter towers while JPT jets (“jet plus tracks”) use additional information from the tracker. Note that only PF jets are used in the measurement of the Z^0 angular distribution.

Residual Calibration

A dijet p_{T}-balance technique, described in [13], is used to measure the response of a jet at any η relative to the jet energy response in the region $|\eta| < 1.3$. Ideally, the relative response of the corrected jets in the simulation should be equal to unity. However, because of a resolution bias effect, the relative response in the simulation is found to deviate from unity. Figure 3.5 shows the residual corrections as a function of η.

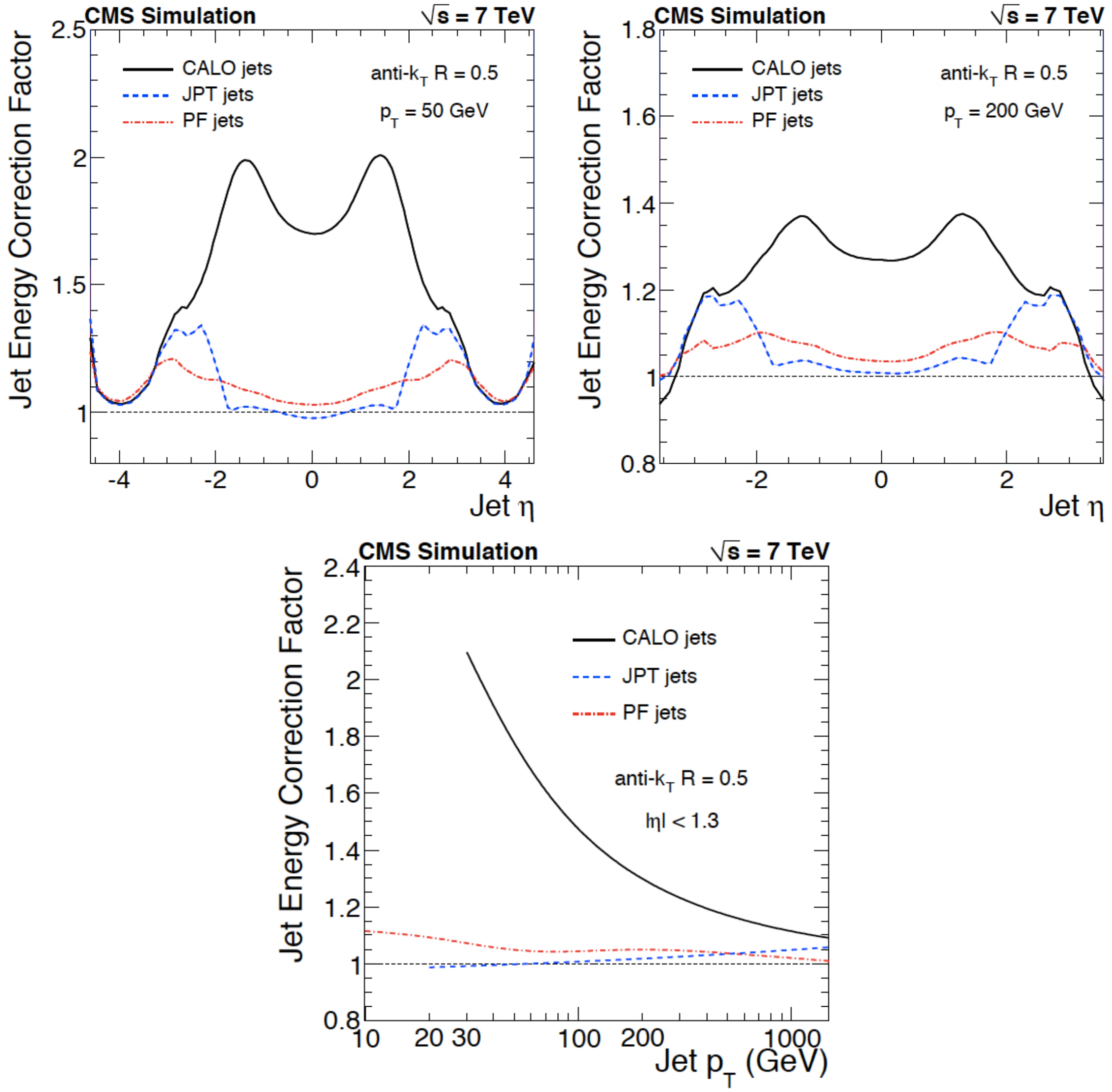


Figure 3.4: Simulation jet energy correction factors for different jet types. The top figures show the correction factors required to get a corrected jet with $p_{\mathrm{T}} = 50$ GeV (left) and $p_{\mathrm{T}} = 200$ GeV (right), as a function of η. The bottom figure shows the average correction in $|\eta| < 1.3$ as a function of jet p_{T} [13].

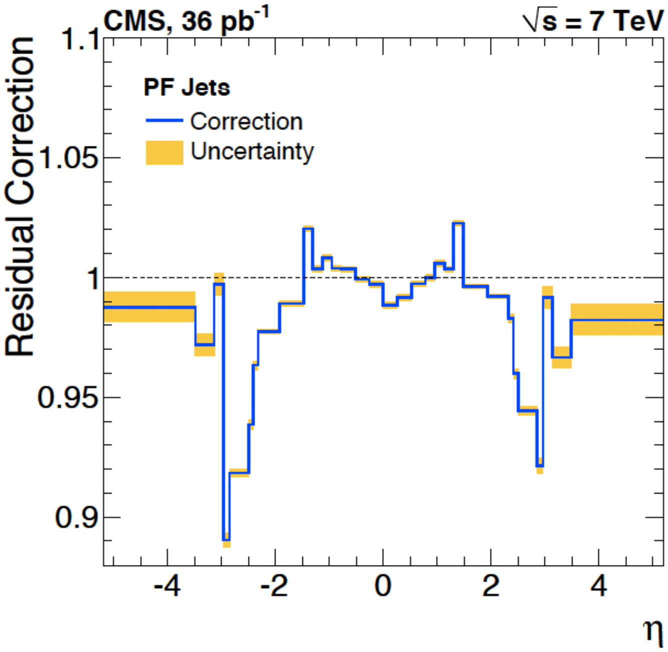


Figure 3.5: Relative jet energy residual correction as a function of η for PF jets. The band shows the uncertainty from statistics, radiation corrections, and asymmetry in η [13].

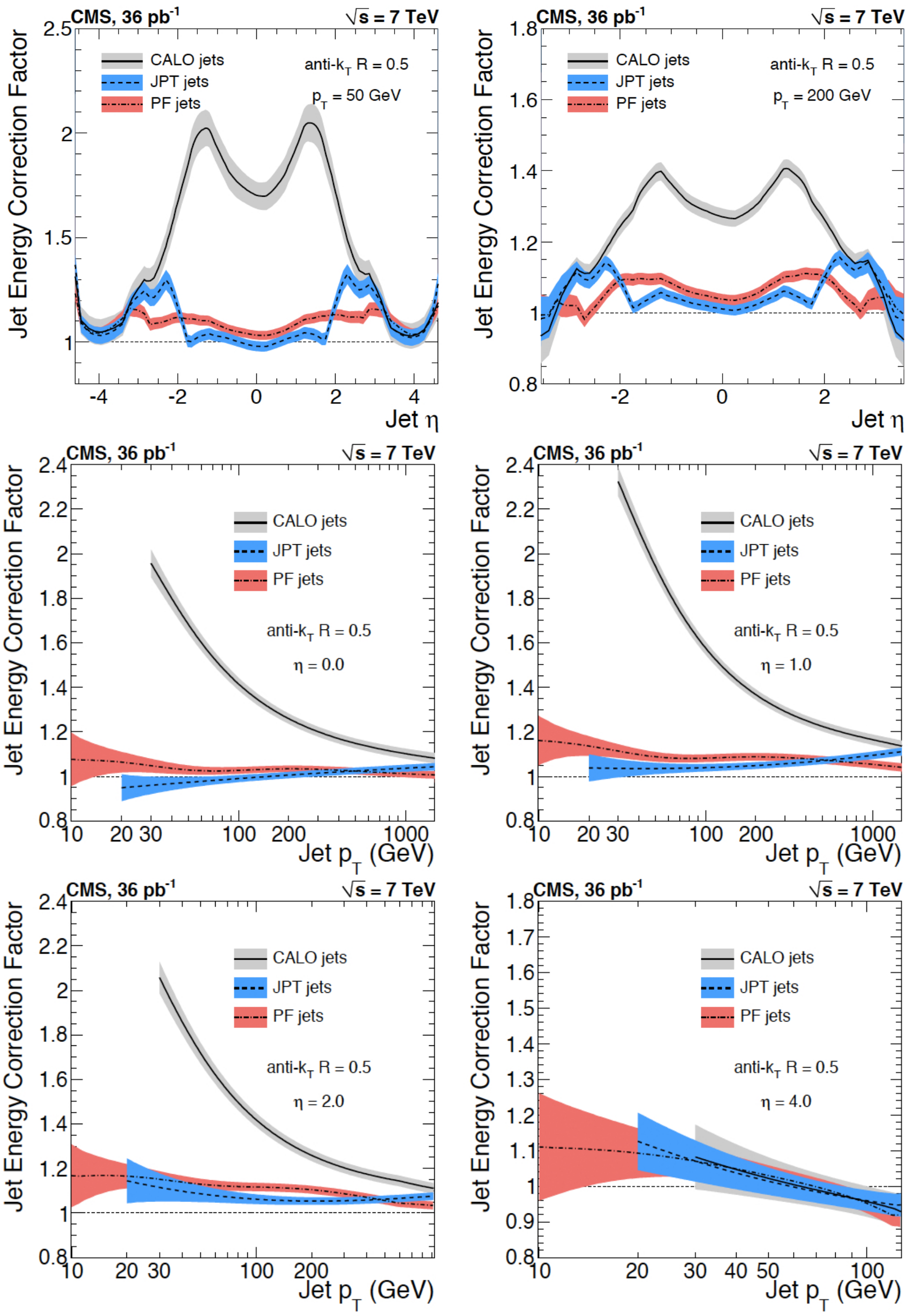


Figure 3.6: Total jet energy correction factor as a function of η for $p_{\mathrm{T}} = 50$ GeV (top-left) and $p_{\mathrm{T}} = 200$ GeV (top-right), and as a function of p_{T} for $\eta = 0$ (middle-left), $\eta = 1$ (middle-right), $\eta = 2$ (bottom-left), and $\eta = 4$ (bottom-right). The bands indicate the corresponding uncertainty [13].

CHAPTER 4

MONTE CARLO SIMULATION

Collision data are compared to QCD theory via Monte Carlo (MC) simulations. Samples of events with a Z^0 boson, $W^{\pm}$ boson, or $t\bar{t}$ pair (accompanied by jets) are generated by MadGraph [20] and interfaced with pythia [21] to simulate parton showering and hadronization. MadGraph generates tree-level events with up to four partons in the final state on the basis of a matrix-element calculation. The pythia parameters for the underlying event have been set according to the Z2 tune, which is identical to the Z1 tune described in [22], except that Z2 uses CTEQ6L [23] PDFs. Various muon-enriched multi-jet events ("QCD backgrounds") are also simulated, although strictly with pythia. Generated events are then processed through a full detector simulation (using geant4 [24, 25]) and reconstructed.

4.1. Matrix Element and Parton Showering

The MadGraph generator is an example of a next-generation MC generator and is better able to represent next-to-leading order and above approximations. To understand why, it is instructive to consider the simulation of multi-jet events. What follows is a brief overview of two commonly used approaches (full descriptions can be found in [20]).

The first involves using the parton model to generate the simplest possible final state (e.g., a $2 \rightarrow 2$ hard process); additional jets are then produced by parton showering (PS). The disadvantage of this approach is that the PS is incapable of describing configurations with several widely separated hard jets and also misses important interference effects. The second approach involves using a tree-level matrix element (ME) with multi-parton final states. The virtue of MEs is that they are exact (up to a given perturbative order) and take all interference effects into account. To avoid soft and collinear divergences at tree-level, ME final state partons have to be well separated in phase space. However, those final state partons have to be converted into hadrons and any realistic model will include PS to connect the hard production scale with the soft hadronization scale. The conversion introduces the possibility of an ME approach producing extra jets or double counting jet configurations after the parton showering.

Although MEs and PS are capable of describing jets in two separate domains, MadGraph uses them in a consistent and systematic way by matching tree-level MEs with PS. The idea is to divide the phase space into two disjoint regions, a jet *production* region filled by the MEs and a jet *evolution* region filled by the PS. The PS then provides a connection between the hard jet production scale and the soft hadronization scale. However, one difficulty of this approach is the existence of several hard scales that span a wide range of energies. The range makes it difficult to separate events that belong to the jet production phase from those that belong to the jet evolution phase. For example, a given 2-jet event can be obtained in two ways: through the *soft and collinear* shower evolution of a 2-parton final state or through a *hard, large angle* radiation in a 1-parton final state. A simplified matching scheme is illustrated in Fig. 4.1. Although the primary goal is to avoid double counting, the matching must also avoid dead regions by ensuring that all configurations are generated by either MEs or PS. Furthermore, physical observables must be independent of the method of phase space separation.

4.2. Pile-Up Simulation

During the high luminosity phase of its operation, the LHC accelerator will produce an average of about 17.5 inelastic *pp* collisions per bunch crossing that will “pile-up” on top of the signal collision firing the trigger. For the MC simulations used in the analysis, diffractive collisions are also considered and increase the pile-up total to 25 collisions [12]. Therefore, several minimum bias events are superimposed to the

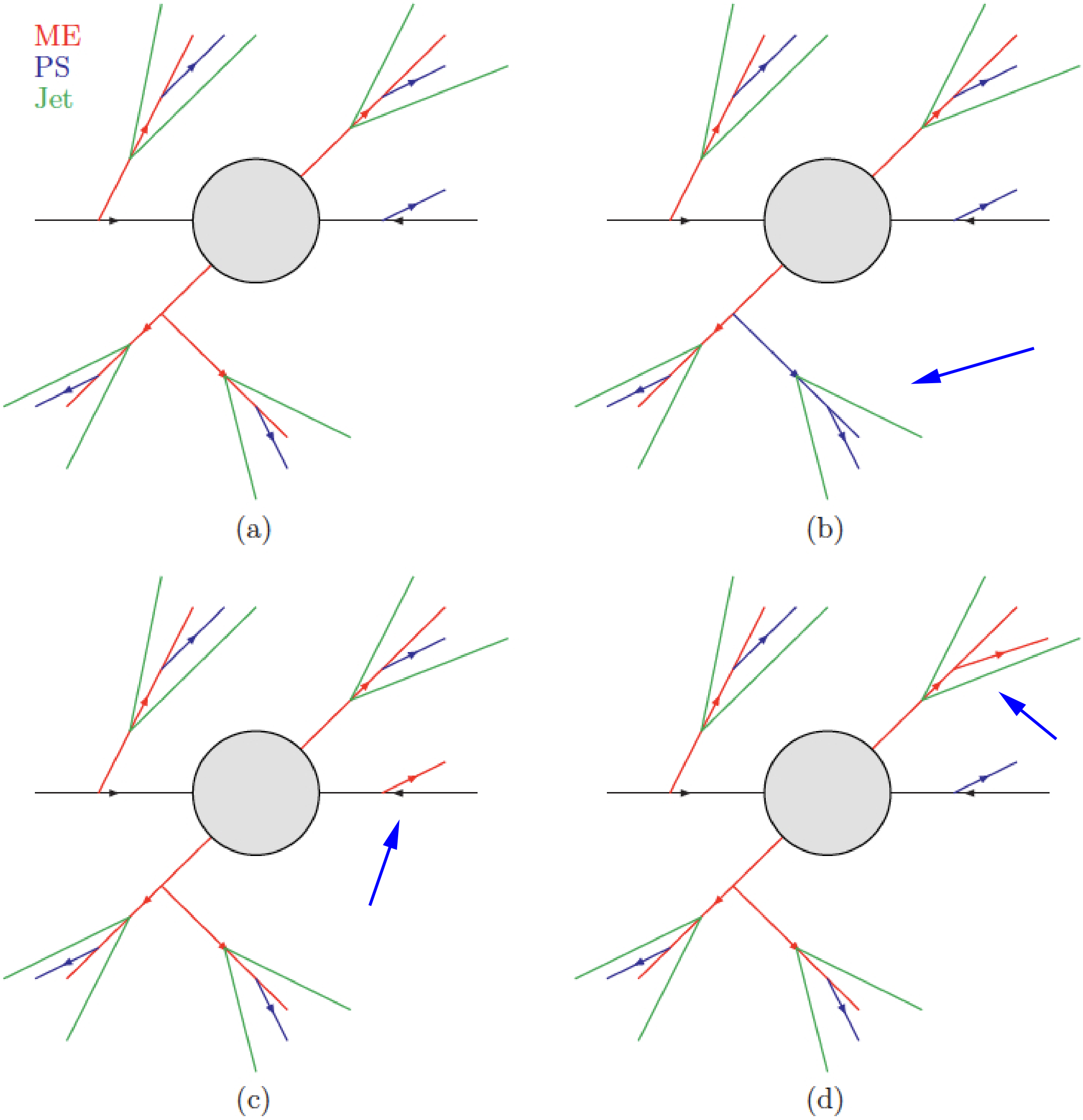


Figure 4.1: Various ME and PS matching examples showing final ME partons (red), PS partons (blue), and jets found via a recombination scheme (green). Although each figure has the same structure, the origin of the partons are different. (a) Accepted - this event is fully matched ($N_{jet} = n_{parton}$). (b) Rejected - this event has one additional hard PS jet ($N_{jet} > n_{parton}$). (c) Rejected - this event has one additional soft ME parton ($N_{jet} < n_{parton}$). (d) Rejected - this event has two ME partons inside the same jet ($N_{jet} < n_{parton}$).

hard interactions to simulate event pile-up according to the distribution of multiple pp collisions observed during the 2010 data taking period.

4.3. Simulation Datasets

The full list of datasets used is given in Table 4.1. Whenever available, the [next-to-]next-to-leading-order ([N]NLO) cross section is used to normalize the simulated sample to the integrated luminosity of the collision data [26, 27]. Scale uncertainties are determined by varying both factorization and renormalization scales by a factor two up and down. The PDF uncertainties are taken as the combination of the variations induced by the 20 × 2 fluctuations in the NLO PDF parameters. These variations were performed by the CMS collaboration.

Table 4.1: Simulation datasets used in the analysis, along with kinematic selections and cross sections. The scale and PDF uncertainties for MadGraph samples are also listed.

Process	Generator	Kinematic Selections	Cross Section (pb)
Z^0 + jets	MadGraph	$m_{\ell\ell} > 50$ GeV	$3048 \pm 34 \pm 128$
$\mathrm{W}^{\pm}$ + jets	MadGraph	–	$31\,314 \pm 407 \pm 1504$
$\mathrm{t\bar{t}}$ + jets	MadGraph	–	$158 \pm 19 \pm 14$
μ-enriched QCD	pythia	$p_{\mathrm{T}}^{boost} > 20$ GeV, $p_{\mathrm{T}}^{\mu} > 5$ GeV, $\lvert\eta_\mu\rvert < 2.5$	3.5×10^6

CHAPTER 5

EVENT IDENTIFICATION

Signal events in collision data are identified through a series of trigger and identification selections. The selections used in the analysis are listed in Table 5.1 and detailed below. Since the goal of the analysis is to study the relative Z^0 angular distribution, the triggers and selections used must not introduce angular biases (e.g., η dependence). A study on the η dependence of our muon selection is discussed in Chapter 7.

Table 5.1: Event and particle selections used in the analysis, along with the motivation.

Category	Selection	Motivation
Vertex	$NDOF_{PV} > 4$ $\|z_{PV}\| < 15$ cm $\rho_{PV} < 2$ cm	Non-collision and beam-related background rejection
Muon Acceptance	$p_{\mathrm{T}}^{\mu_1} > 20$ GeV $p_{\mathrm{T}}^{\mu_2} > 10$ GeV $\|\eta_\mu\| < 2.1$	... ... Muon trigger acceptance
Muon Quality	Global Muon & Tracker Muon Number of Pixel Hits > 0 Number of Silicon Hits > 10 Number of Muon Hits > 0 Number of Muon Stations > 1 Normalized $\chi^2 < 10$ $\|d_{xy}\| < 0.2$ cm	Decay-in-flight, punch-through, noise and cosmic ray rejection
Jet Acceptance	$p_{\mathrm{T}}^{jet} > 20$ GeV $\|\eta_{jet}\| < 2.4$	... Tracker and muon spectrometer acceptance
Z Selection	$I_{rel}^{\mu} < 15\%$ $60 < M_{\mu\mu} < 120$ GeV	QCD background rejection Z^0 signal selection

5.1. Trigger Selection

For the analysis, events were selected if they passed any of the muon triggers listed in Table 5.2. The rates of various triggers were prescaled since the instantaneous luminosity increased throughout the 2010 run. Therefore, events were accepted only if they passed an unprescaled trigger.

Table 5.2: Trigger paths used in selecting events including HLT and L1 p_{T} thresholds.

HLT Path	L1 Seed	HLT/L1 p_{T} Threshold (GeV)
HLT_Mu9	L1_SingleMu7	9 / 7
HLT_Mu11		11 / 7
HLT_Mu13		13 / 7
HLT_Mu15		15 / 7

5.2. Collision Selection

To reject non-collision and beam-related backgrounds, all events are required to have a primary vertex (PV) consistent with the measured transverse position of the beam (referred to as the beam spot). Specifically, the fit for the PV must include at least four associated tracks (i.e., five degrees of freedom), the z-coordinate of the PV must lie within the luminous collision region, and the radial distance of the PV must be less than 2 cm from the beam spot. These selections are greater than 99% efficient; the full studies are discussed in [28].

5.3. Muon Identification

Our signal muon selections are those used in the measurement of the W and Z^0 cross sections [29], with isolation modifications motivated by the vector boson + jet ratio measurements [30]. The modified definition of isolation is found to be more effective at suppressing the multi-jet background. The relevant selections are summarized below and are all greater than 99% efficient.

We define a relative isolation variable $I_{rel} = \sum \left(p_{\mathrm{T}}^{track} + E_{\mathrm{T}}^{ECAL} + E_{\mathrm{T}}^{HCAL} \right) / p_{\mathrm{T}}^{\mu}$, which consists of the p_{T} for tracks and E_{T} for calorimeter towers within a cone of $R < 0.3$ centered around the muon's trajectory. The muon and its energy deposits are excluded from this sum by ignoring the energy within a smaller "veto cone" ($R < 0.01$, 0.7 and 0.1 for the tracker, ECAL and HCAL, respectively). A muon is considered isolated if the energy within its isolation cone is less than 15% of the muon p_{T}, i.e., $I_{rel} < 0.15$ [30].

5.4. Z+Jet Identification

The Z+jet event selection requires the presence of an energetic ($p_{\mathrm{T}} > 20$ GeV), isolated muon in the region $|\eta| < 2.1$. We then require a second muon ($p_{\mathrm{T}} > 10$ GeV within $|\eta| < 2.1$) such that the dimuon invariant mass is between 60 and 120 GeV. The event must also contain at least one jet with a p_{T} of more than 20 GeV within the muon system and tracker acceptance ($|\eta| < 2.4$).

5.5. Kinematic Distributions of Candidate Events

Figures 5.1 and 5.2 show kinematic distributions for candidate Z^0s and jets, respectively. After selection, all distributions agree with simulations within statistical and systematic uncertainty. Note that the jet mass in Fig. 5.2 is non-zero; this is attributed to the finite angular spread caused by hadronization. The Z^0 mass distribution shown in Fig. 5.1 was created before applying the Z^0 mass selection; the discrepancy in collision data and simulation (for $M < 50$ GeV) comes from a generator-level invariant mass selection (see Table 4.1).

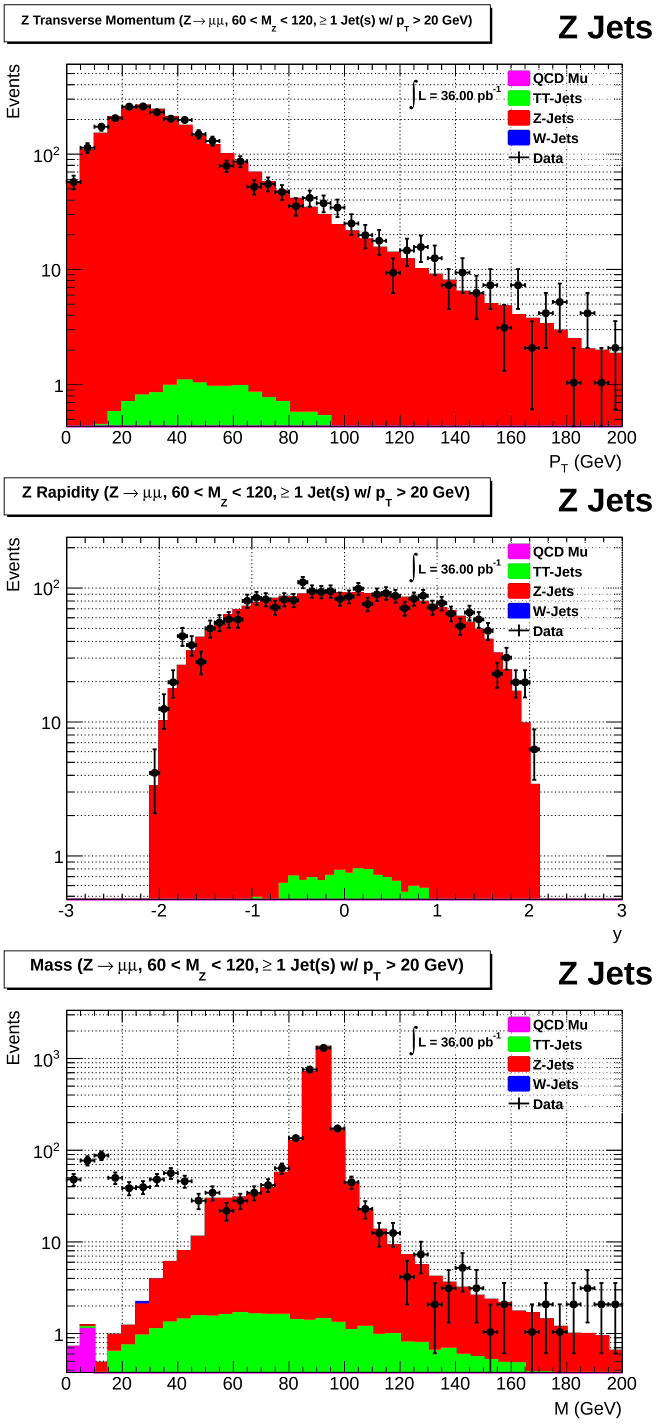


Figure 5.1: Kinematic distributions for candidate Z^0s in collision data; transverse momentum (p_{T}, top), rapidity (y, middle), and mass (M, bottom). Also shown are the corresponding distributions for signal (red) and various background (other) simulations.

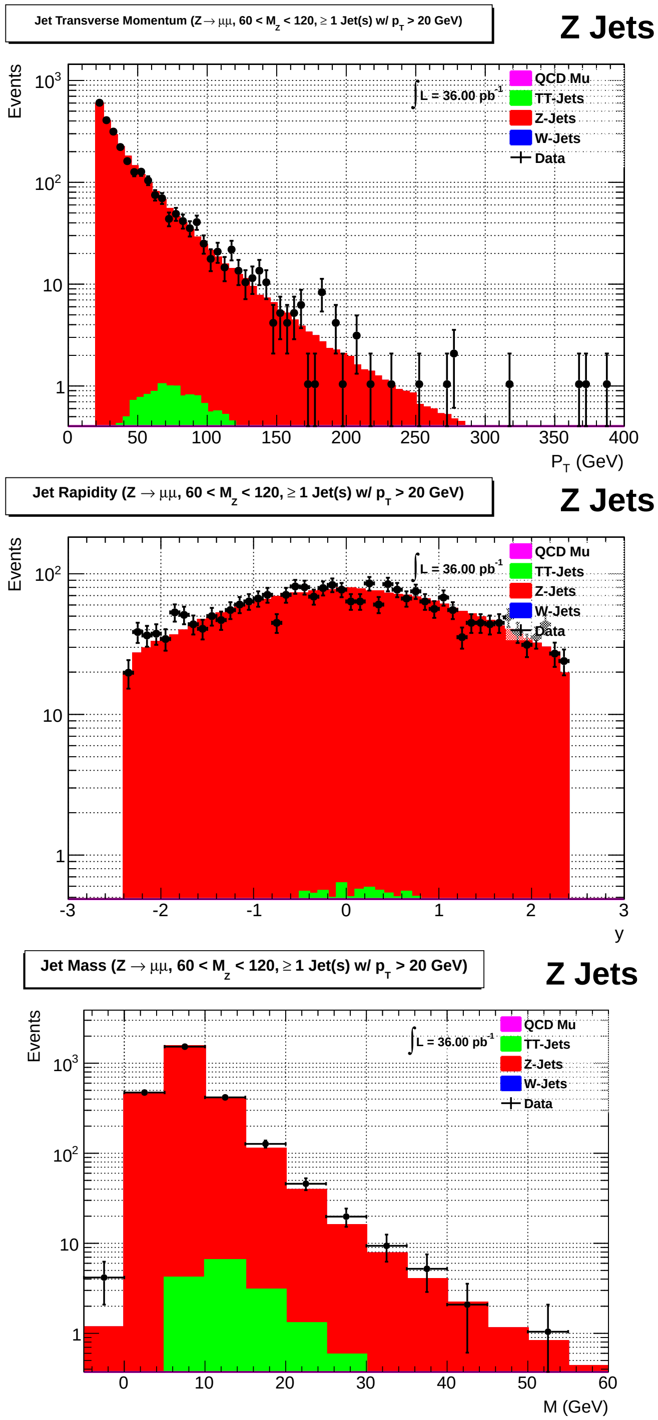


Figure 5.2: Kinematic distributions for candidate jets in collision data; transverse momentum (p_{T}, top), rapidity (y, middle), and mass (M, bottom). Also shown are the corresponding distributions for signal (red) and various background (other) simulations.

CHAPTER 6

CENTER OF MOMENTUM KINEMATICS

The purpose of the analysis is the measurement of the Z^0 angular distribution in the CM frame. We must therefore reconstruct the CM kinematic variables from measurements taken in the lab frame. The boosted system (see Fig. 1.4) can be reconstructed by summing the lab frame four-momentum vectors of the Z^0 and jet. However, we must first decide how to treat the four-momentum vectors of multiple jets in an event.

As mentioned in Chapter 1, QCD jets can arise from the fragmentation of a hard-scatter parton (LO) and/or from the initial- and final-state radiation of a quark or gluon (NLO and above). For the analysis, we are only interested in the scattering between the parton and the Z^0. However, it is not possible to distinguish between the above two scenarios since radiation can be energetic and emitted in all directions. Thus there are three possible methods to use when considering Z+jet events:

- *Single-jet*: Consider only Z^0 events with one jet.
- *Leading-jet*: Consider all Z^0 events but only use the leading p_{T} jet in kinematic calculations.
- *Multi-jet*: Consider all Z^0 events and use the sum of all jet four-momentum vectors in kinematic calculations.

Initially one would think that the single-jet method is preferable since we are studying the scattering of a Z^0 and a single parton. Nevertheless, this leads to a limited amount of candidate Z+jet events. Figure 6.1 shows the number of jets accompanying a Z^0 (note that adding a jet drops the number of events by $\sim \alpha_s$; the cross section is proportional to the number of strong-interaction vertices in the Feynman diagram). The figure shows that $\approx 10\%$ of events have two or more jets.

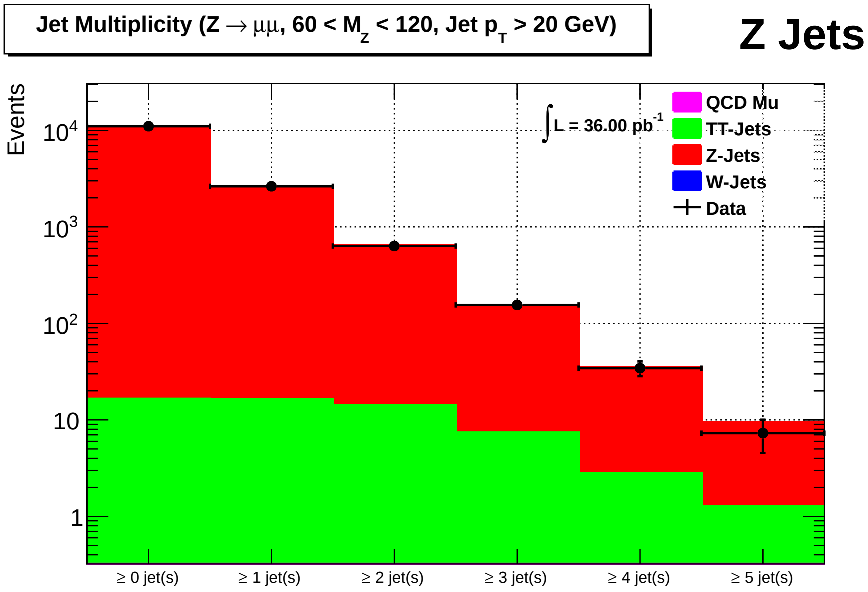


Figure 6.1: The number of jets accompanying Z^0 events. Collision data (marker) and signal simulations (red) are shown. Background simulations are also shown, but are negligible.

To motivate a choice between methods, we must go back to the initial collision. It is assumed that the collision of the two incoming partons takes place strictly along the z-axis; at LO, the system has no initial or final total p_{T}. Therefore, the azimuthal angle between the Z^0 and the jet ($\Delta\phi$) tells us how well the jet balances the p_{T} of the Z^0. Particularly, a well-defined Z+jet event will be back-to-back in ϕ.

The $\Delta\phi$ distribution is shown in Fig. 6.2 and gives similar results for all methods. We conclude that the leading jet does as sufficient a job of balancing the Z^0 in events with multiple jets as in events with only one jet. Consequently, in order to minimize the loss of data, we do not use the single-jet method. Interestingly, a choice between the leading- or multi-jet method cannot be justified. The multi-jet method intuitively handles radiation as it would recombine radiated jets with their parent, but it does not make sense to combine multiple hard-scatter jets (i.e., partons from the ME) into one jet. Thus we show the angular distribution for both in Chapter 8. For simplicity, we focus on the leading-jet method throughout this dissertation and further compare the two in Appendix C.

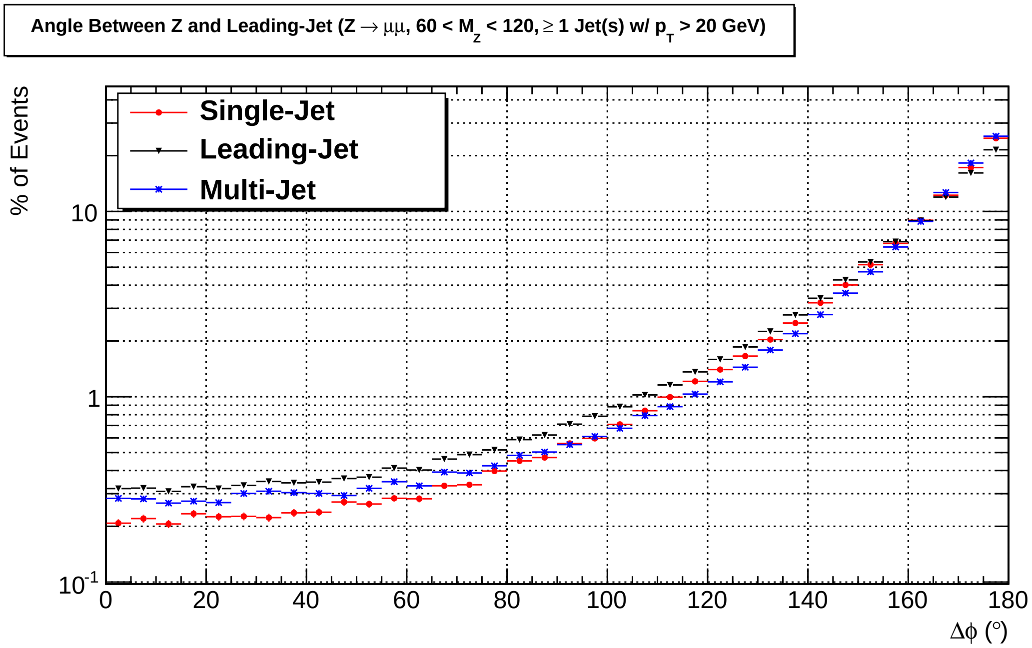


Figure 6.2: The azimuthal angle between the Z^0 and jet using three methods in simulated signal; single-jet (red circle), leading-jet (black triangle) and multi-jet (blue star).

We can now reconstruct the boosted system (by summing the four-momentum vectors of the Z and jet in the lab) and use it to perform a Lorentz boost, transforming the four-momentum vector of the Z^0/jet from the lab frame to the CM frame.

Note that, to first order, the initial partons in the pp collision have zero p_T; this implies that the boosted system will also have zero p_T. Nevertheless, Fig. 6.3 shows that the boosted system has non-zero p_T. It has been shown that the intrinsic transverse momenta (k_T) of the initial state partons can be a few hundred MeV as a consequence of the finite size of the proton [31]. In addition, initial state soft gluon radiation (a NLO effect) can generate sizable transverse components of the parton momenta. Furthermore, some uncertainty is introduced in the total p_{T} of an event because of the loss of particles/energy down the beam pipe.

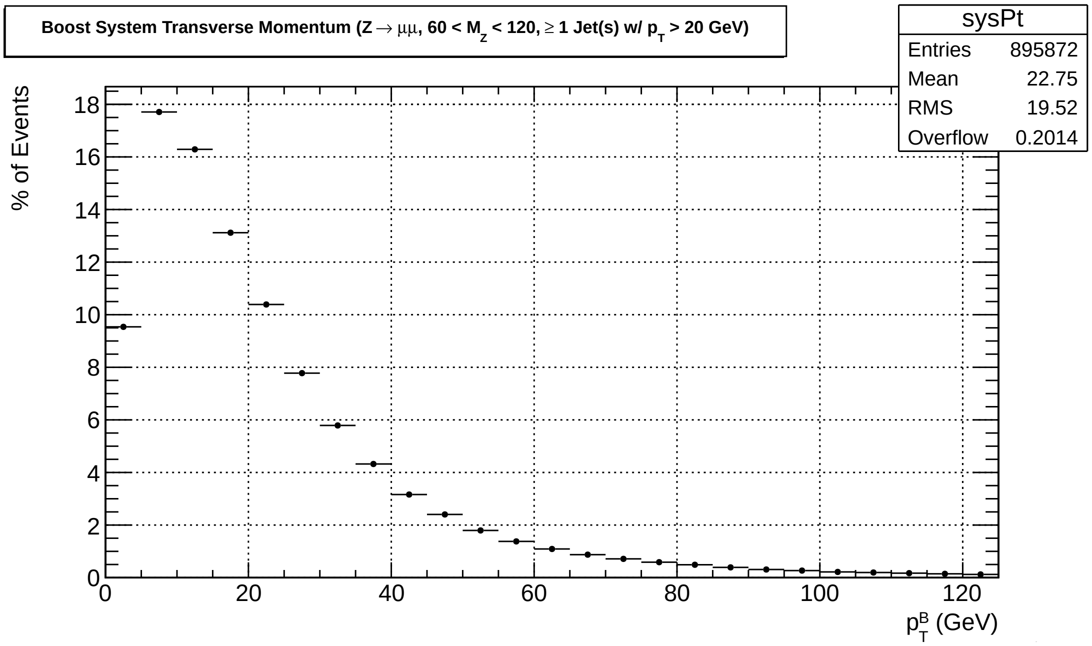


Figure 6.3: The boosted system's p_{T} in simulated Z+jet signal.

It is informative to study the relationship between lab and CM frame variables explicitly. Since rapidities are additive under Lorentz transformations, one can write [4]

$$y = y_B + y^*, \tag{6.1}$$

where y^* and y are the rapidities of the Z^0/jet in the CM and lab frame, respectively, and y_B is the rapidity of the boosted system. One can also write the energy, E^*, and longitudinal momentum, p_z^*, as a function of rapidity,

$$\begin{aligned} E^* &= m_{\mathrm{T}} \cosh y^* \\ p_z^* &= m_{\mathrm{T}} \sinh y^* \\ m_{\mathrm{T}} &\equiv \sqrt{M^2 + p_{\mathrm{T}}^2}, \end{aligned} \tag{6.2}$$

where m_T is the transverse mass and M is the mass of the particle ($M = 0$ for jets). Equations 6.2 hold true in all reference frames.

One can then derive the angular distribution by calculating $p_z^*/E^* = \tanh y^*$ and noting that p_z^*/E^* can be rewritten as $p^* \cos\theta^*/E^* = \beta^* \cos\theta^*$ (since $p^*/E^* = \beta^*$). Therefore, $p_z^*/E^* = \beta^* \cos\theta^*$. Finally, equating both ratios gives

$$\beta^* \cos\theta^* = \tanh y^*, \tag{6.3}$$

which is the angular distribution of the Z^0 in the CM frame.

CHAPTER 7

CORRECTIONS AND UNCERTAINTIES

7.1. Corrections

The relative Z^0 angular distribution was measured, so it is not necessary to correct for absolute detector acceptance or trigger/selection efficiency effects. We must however investigate whether the selections we perform introduce an angular bias (e.g., an η dependence).

7.1.1. Single Muon Efficiency

A selection dependence on η is indeed observed (see Fig. 7.1) because separate η regions are covered by muon detectors with a variety of structural material. The dependence is confirmed and further discussed in [29]. Fortunately, a bin-by-bin correction of the η inefficiencies has a negligible effect on the angular distribution (see Appendix B) and is therefore not applied. However, to further complicate matters, the dependence is not well simulated and the discrepancy must be corrected in the MC.

Figure 7.2 shows the ratio

$$\rho = \frac{\epsilon_{\text{data}}}{\epsilon_{\text{sim}}}, \tag{7.1}$$

where ϵ_{data} is the efficiency in collision data and ϵ_{sim} is the efficiency in simulation. These efficiencies were calculated in [29] using a “tag and probe” method. In short, the method requires a mass resonance (e.g., Z^0 candidate event) that is used to compare specific selection criteria on both leptons. One lepton candidate (called the tag) must satisfy all selections (and have a low fake rate). The other lepton candidate (called the probe) is required to pass the specific criteria whose efficiency is under study. The efficiency is then the number of passing probes divided by the total number of probes.

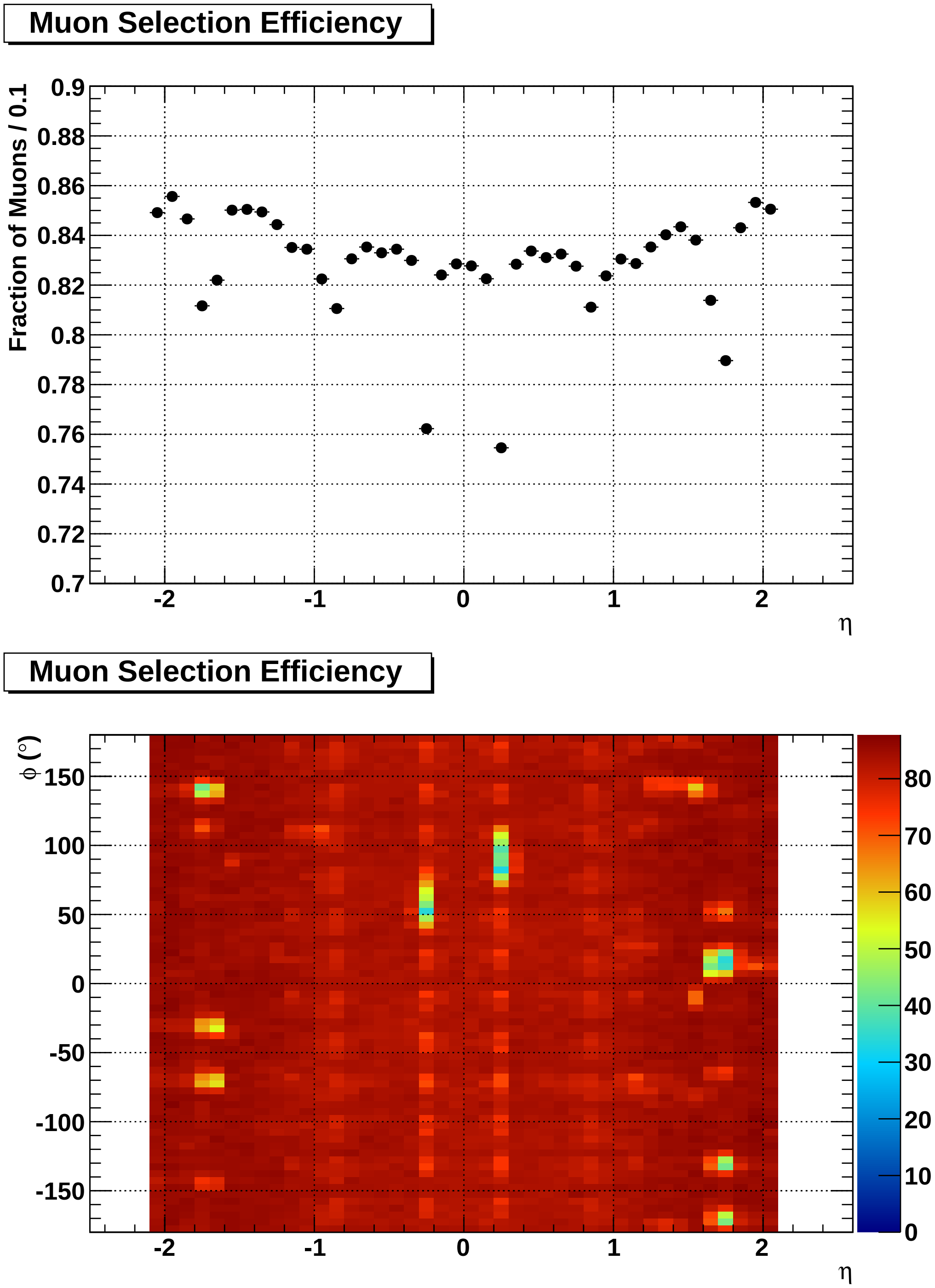


Figure 7.1: Single muon selection efficiency as a function of η (top) and η,ϕ (bottom), for a simulated Z+jet signal.

The discrepancy was parameterized as $\rho = 0.99 - 0.018\,\eta^2$ and used to discard a random fraction $(1-\rho)$ of the MC events (see Appendix B).

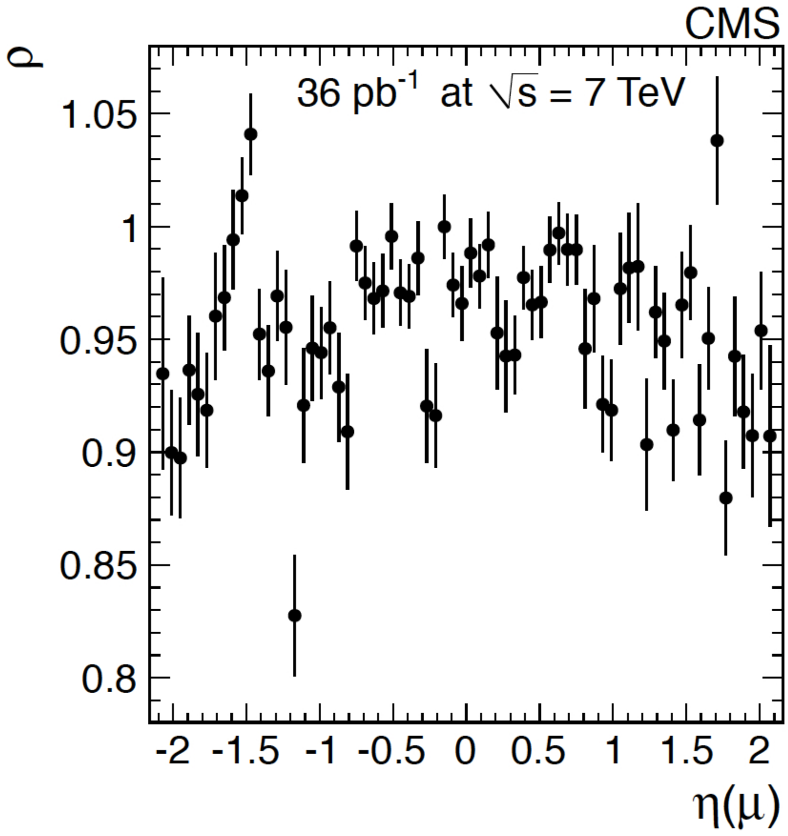


Figure 7.2: Ratio of collision data to simulation single muon η efficiency. The points at $|\eta| \approx 1.6$ are at the interface between the RPCs and CSCs (see Fig. 2.11) [29].

7.1.2. Phase Space Bias

Another bias comes from selections applied in the lab frame. Muons and jets are constrained to the central region of the detector ($|\eta| < 2.4$) because identification/resolution in the forward region is not sufficiently understood. This ultimately limits the reach of the $\cos\theta^*$ measurement (Eq. 6.3). Although a more significant effect is the p_{T} selection of the jet; it introduces a bias in the data through the relationship between p^* and y^*.

Since $p^* = p_\mathrm{T} \cosh y^*$ (Eq. 6.2 in the relativistic limit), a fixed lower value of p_T results in a loss of acceptance that increases with y^* (see Fig. 7.3). The effect is a flattening of the $\cos\theta^*$ distribution. Unfortunately, it is the high values of $\cos\theta^*$ that are of interest in pQCD; additional radiation moderates the LO collinear singularities at $\cos\theta^* = \pm 1$ (recall that LO predicts the angular distribution to be of the form $(1 - |\cos\theta^*|)^{-1}$). In order to reduce the CM phase space bias we must set a limit on the p^* integration boundary of the invariant cross section. The minimum value of p^* will inversely affect the amount of statistics available and the $\cos\theta^*$ reach. Therefore, a choice can be motivated by studying Eq. 6.2;

$$\begin{aligned} p^*_\mathrm{min} &= p_\mathrm{T,min} \cosh y^*_\mathrm{max} \\ y^*_\mathrm{max} &= \cosh^{-1}\left(\frac{p^*_\mathrm{min}}{p_\mathrm{T,min}}\right), \end{aligned} \tag{7.2}$$

but Eq. 6.3 states (for $\beta_\mathrm{jet} = 1$) that

$$\cos\theta^*_\mathrm{max} = \tanh y^*_\mathrm{max} = \tanh\left[\cosh^{-1}\left(\frac{p^*_\mathrm{min}}{p_\mathrm{T,min}}\right)\right]. \tag{7.3}$$

The relationship between p^*_min and $\cos\theta^*_\mathrm{max}$ is shown in Fig. 7.4 (for $p_\mathrm{T,min} = 20$ GeV). Note that this relationship is smeared by the finite detector resolution (particularly, the p_T measurement in the CM is smeared because of NLO and measurement effects – see Chapter 6); this is why we see events below the acceptance limit in Fig. 7.3. To minimize the loss of statistics we choose a value of $p^*_\mathrm{min} = 45$ GeV; this results in a $\cos\theta^*$ reach of approximately 0.90 – higher values of the measured angular distribution will affected by the acceptance. The resulting p^* distribution is shown in Fig. 7.5.

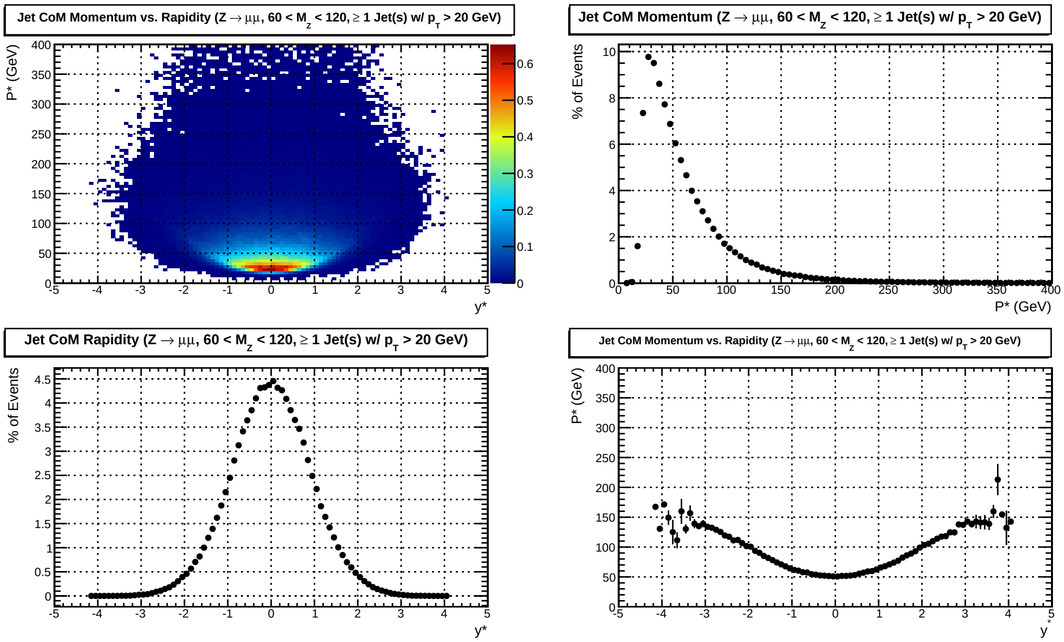


Figure 7.3: CM energy *vs.* rapidity for the jet in simulated Z+jet signal. The top-left is a 2D distribution of P^* *vs.* y^*; the top-right and bottom-left are 1D distributions of P^* and y^*, respectively; and the bottom-right is the 2D profile.

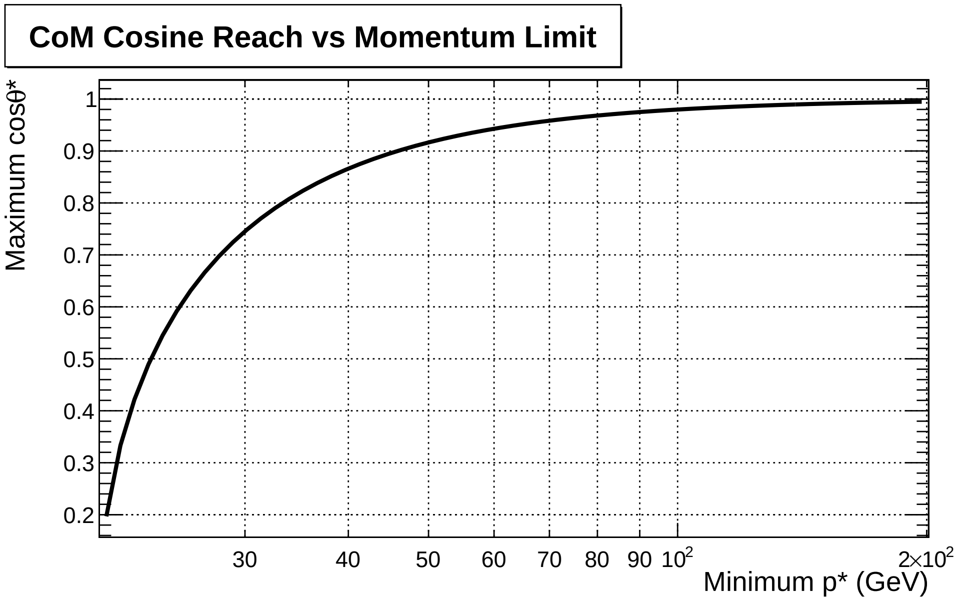


Figure 7.4: CM cosine reach *vs.* momentum for the jet in simulated Z+jet signal.

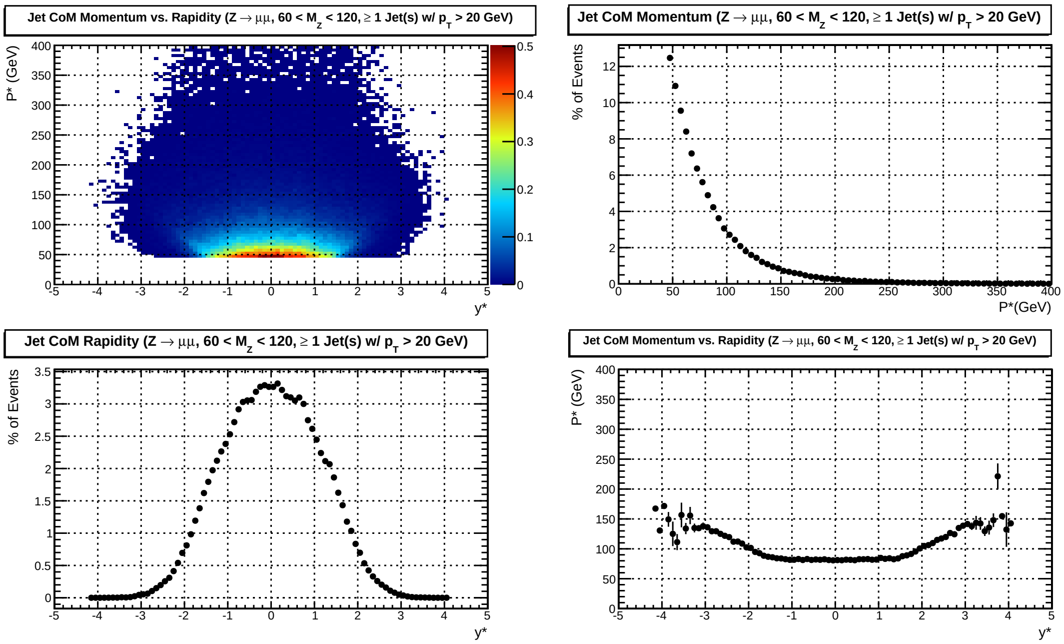


Figure 7.5: CM energy *vs.* rapidity for the jet in simulated Z+jet signal, after correcting for phase space bias. The top-left is a 2D distribution of P^* *vs.* y^*; the top-right and bottom-left are 1D distributions of P^* and y^*, respectively; and the bottom-right is the 2D profile.

7.2. Uncertainties

For a relative angular distribution analysis, many of the typical uncertainty studies are not required. Of the usual sources of uncertainty, we consider only those that affect the shape of the angular distribution (or that have an angular bias). We evaluate how varying the PDF parameters in our simulation affects the shape of the angular distribution. We also study the effects of uncertainty in the JES and take the finite jet momentum and angular resolution of the detector into account. Compared to the JES and resolution, the uncertainty in the muon measurement is negligible (see Chapter 2). Relative uncertainty values are summarized in Table 7.1.

7.2.1. PDF Uncertainty

Crucial to higher order cross section predictions, PDFs are obtained by global fits to measurements from deep-inelastic scattering, Drell-Yan, and jet data. The PDF uncertainties reflect three characteristics: the choice of dataset, the type of uncertainty estimator used, and the form and size of parton parametrization. Details can be found in [32], with a brief summary below.

One method of determining PDFs is based on a Hessian approach, which minimizes a suitable log-likelihood χ^2 function and accounts for correlated uncertainties by means of a covariance matrix. The best fit is the point in parameter space at which χ^2 is minimum. The PDF uncertainties are found by diagonalizing the Hessian matrix (second derivatives of the χ^2 at the minimum) and then determining the range of each orthonormal Hessian eigenvector that corresponds to a prescribed increase (e.g., a 68% confidence level) of the χ^2 function with respect to the minimum.

The above method was applied to CTEQ10 [33] PDFs on a photon+jet cross section [34], with the results shown in Fig. 7.6. Although the cross section is different, the shape (and the PDF uncertainty's effect on the shape) is analogous to Z+jets. Note that the Hessian uncertainty is inflated by a lack of MC statistics, we therefore choose a conservative relative PDF uncertainty of 5%. This method is similar to that applied by the MSTW PDF authors [35].

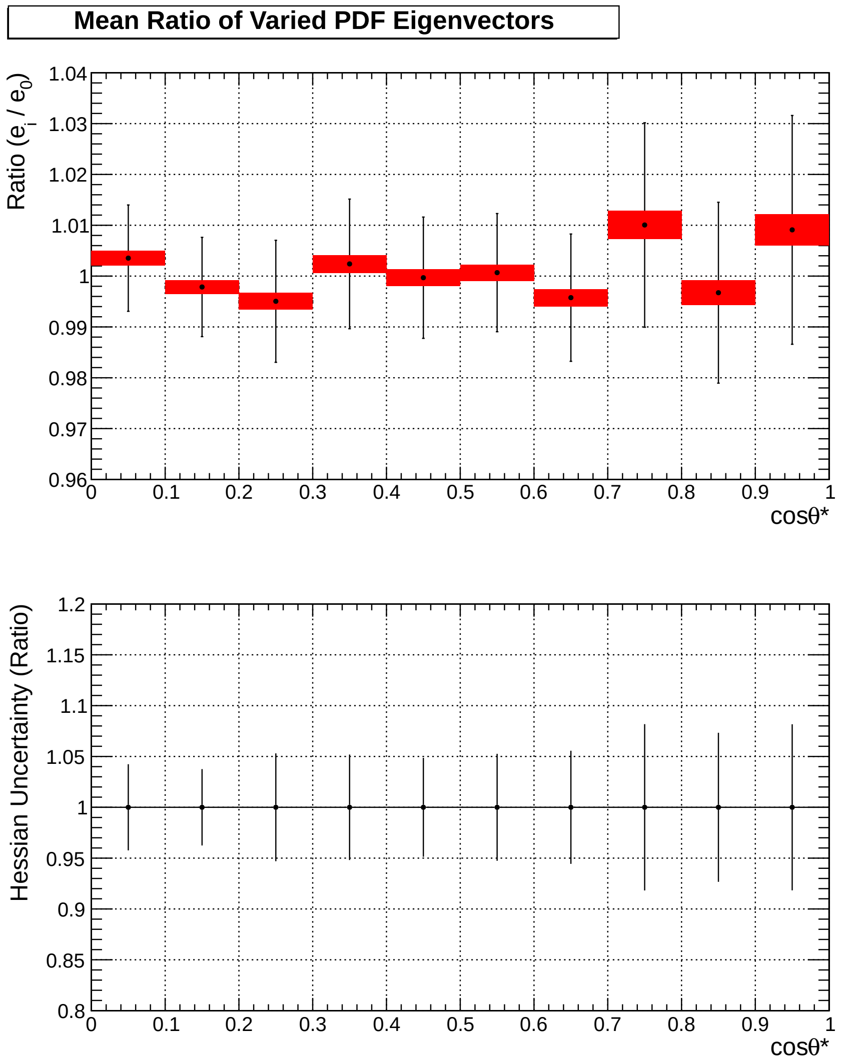


Figure 7.6: PDF Hessian uncertainty. The top plot is the ratio of Hessian eigenvectors to the minimum χ^2 value. The larger (black) error bars represent the spread in the differences while the smaller (red) error bars represent the error in the mean. The bottom plot shows the Hessian uncertainty.

7.2.2. JES Uncertainty

As mentioned in Chapter 3, the measured jet energy is different from the true particle-level jet energy. The difference is primarily caused by the non-uniform and non-linear response of the calorimeters. Therefore, corrections are made to the energy scale of the reconstructed jets. Uncertainty studies on the JES are described in [13] and summarized below.

For the offset correction, the p_{T} density was varied independently and the resulting shifts were added in quadrature. The offset uncertainty was estimated to be 0.2 GeV per unit jet area and per pile-up event. For the relative correction, the comparison of collision data with MC simulation implicitly assumes that the resolution in the data is the same as in the simulation; this assumption is the dominant systematic uncertainty. Consequently, an estimate of the relative uncertainty was achieved by varying the simulated jet p_{T} resolution and was evaluated to be less than 2%. The uncertainty of the absolute JES measurement has various components, which include the photon energy scale, MC extrapolations, and the offset attributed to noise and pile-up. After independently varying each parameter, the relative uncertainty for the absolute correction was found to be between approximately 2% (for high p_{T} jets) and up to 8% (for low p_{T} jets).

The above uncertainties were added in quadrature to obtain the combined JES uncertainty as a function of p_{T} for various η values (shown in Fig. 7.7). However, the JES correction and uncertainty studies have since been updated and new values are stored in an offline database. The uncertainty values corresponding to the processing of the 2010 data are shown in Fig. 7.8. To determine the effect on the angular distribution, we scale the jet energy up and down by its maximum uncertainty value and recalculate $\cos\theta^*$; the results are shown in Fig 7.9. In general, the effect of the JES uncertainty can be evaluated as a relative uncertainty of less than 5% (however, these values are used on a bin-by-bin basis in the final measurement).

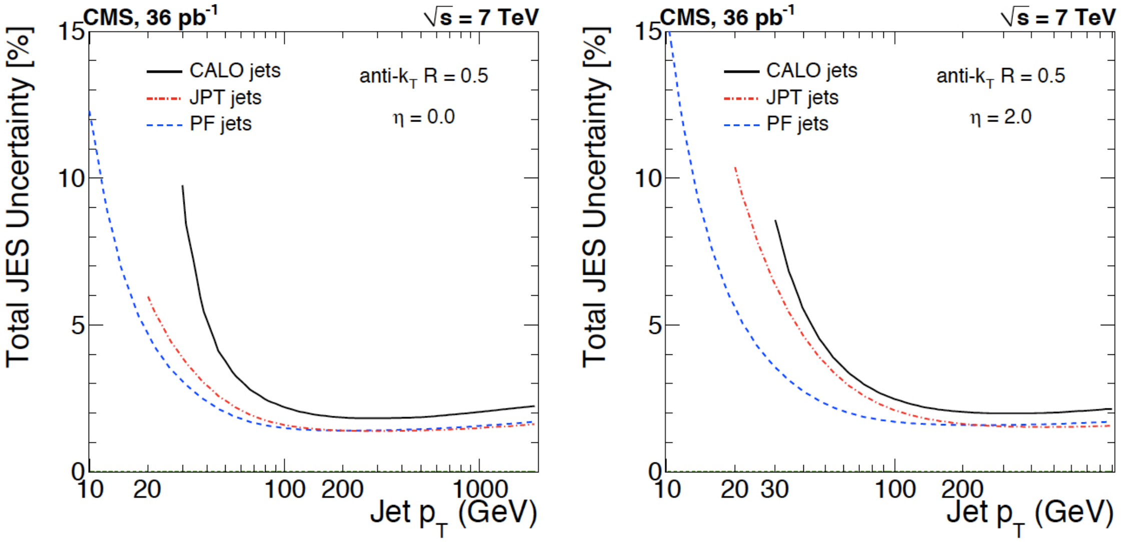


Figure 7.7: Total jet energy scale uncertainty as a function of jet p_{T} for various η values [13].

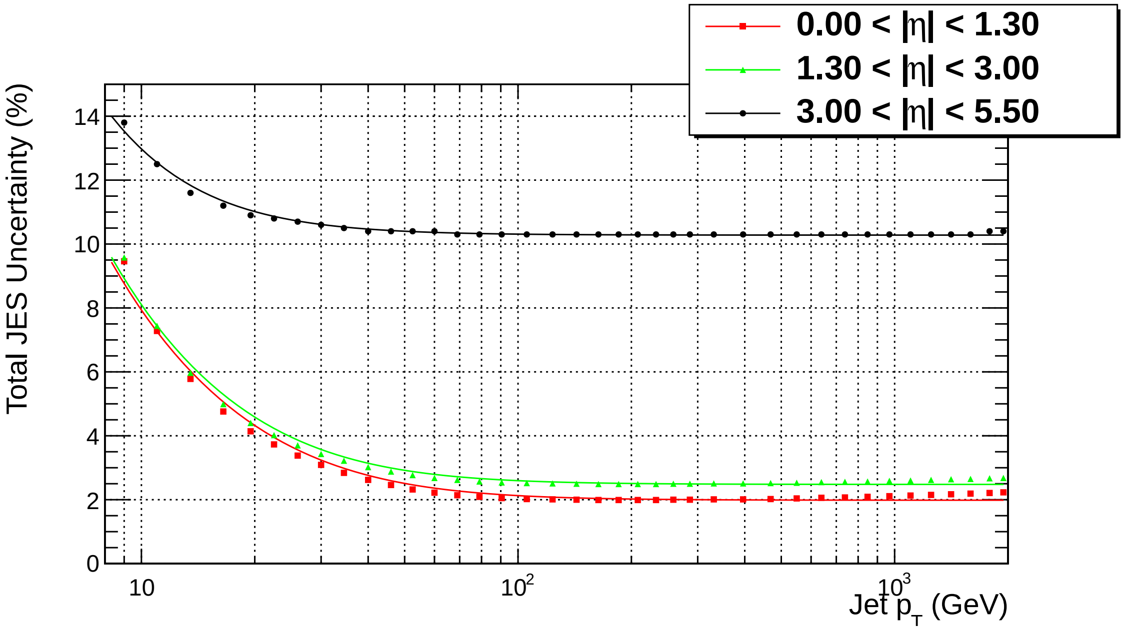


Figure 7.8: Updated combined jet energy scale uncertainty as a function of jet p_{T} for various η values. Note that jets used in the analysis of the angular distribution have $|\eta| < 2.4$.

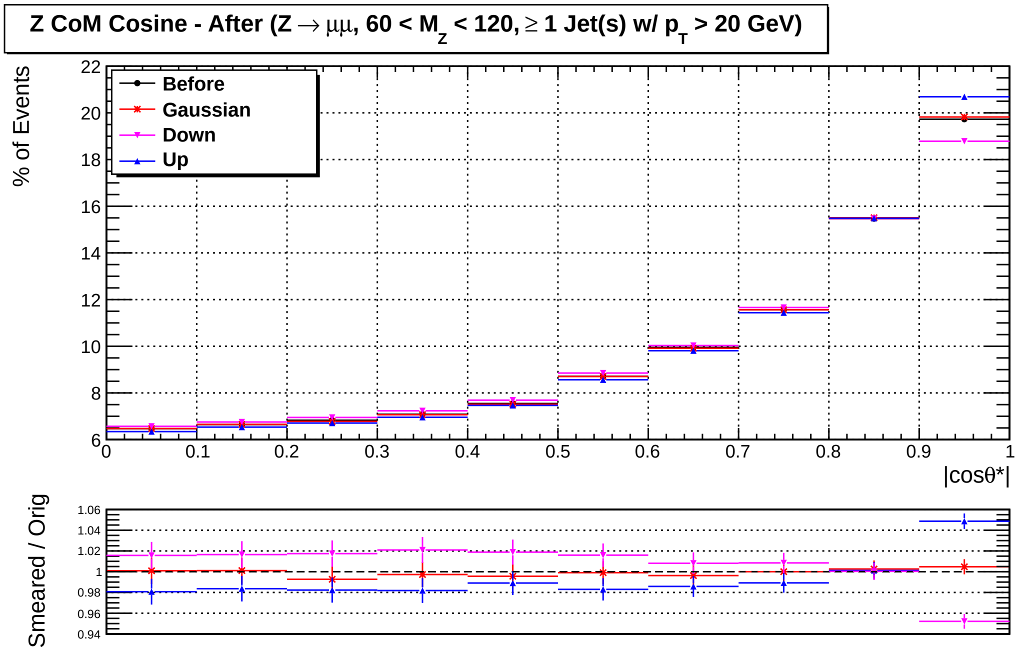


Figure 7.9: The relative uncertainty attributed to JES variations by scaling the jet energy randomly (red star), up (blue up arrow) and down (magenta down arrow) for simulated Z+jet signal.

7.2.3. Jet Resolution

The importance of a particle physics measurement is more clearly understood after comparing it to a given theoretical prediction. Complications arise when a measured value is subject to random fluctuations caused by a finite measurement resolution. Each observation is then characterized by a true (and unknown) value t that is smeared by detector effects and becomes a measured value m. In general, one can simply smear the prediction to include the distortions of the detector. However, the measurement cannot be subsequently compared with the results of other experiments because the smearing is detector dependent. Therefore, the measurement is typically "unfolded" of detector effects. For the analysis, the jet p_{T} and η resolutions have large effects and thus the angular distribution measurement must be unfolded.

The continuous distributions for t and m are related by a convolution [36],

$$f_{\text{meas}}(m) = \int R\left(m|t\right) f_{\text{true}}(t) dt, \tag{7.4}$$

where R is called the response function and depends only on the measuring apparatus. For a measurement binned into a histogram, Eq. 7.4 becomes

$$m_i = \sum_{j=1}^{N} R_{ij} t_j, \tag{7.5}$$

where N is the total number of bins. The response matrix then has the simple interpretation of a conditional probability: R_{ij} is the probability that an observed value in bin i corresponds to a true value in bin j. To obtain the true distribution, one may invert the response matrix such that

$$\mathbf{t} = R^{-1}\mathbf{m}, \tag{7.6}$$

where $\mathbf{t}$ and $\mathbf{m}$ are vectors of the true and measured values, respectively. However, unfolding techniques use regulated procedures (discussed in [36]) in order to minimize the sensitivity to statistical fluctuations. The unfolding procedures in the analysis were performed with the RooUnfold package.

The RooUnfold package [37] provides a common framework to evaluate various unfolding algorithms, providing implementations for the Iterative Bayes [38], Singular Value Decomposition [39], and TUnfold methods [40], as well as bin-by-bin correction factor and unregularized matrix inversion methods. We choose to perform a Bayes unfolding and use the bin-by-bin and matrix inversion methods as references (see Appendix A).

First, an instructive distribution to study is the response matrix shown in Fig 7.10; i.e., the comparison between true (generated) and measured (reconstructed) values. Generated values are taken from the MC generator, before going through a detector simulation, while reconstructed values are taken after a simulation of detector resolution effects. There are minimal off-diagonal elements; this is attributed to the fact that the particle-flow algorithm and JES perform exceptionally well in accounting for detector resolution and in representing the true particle energy.

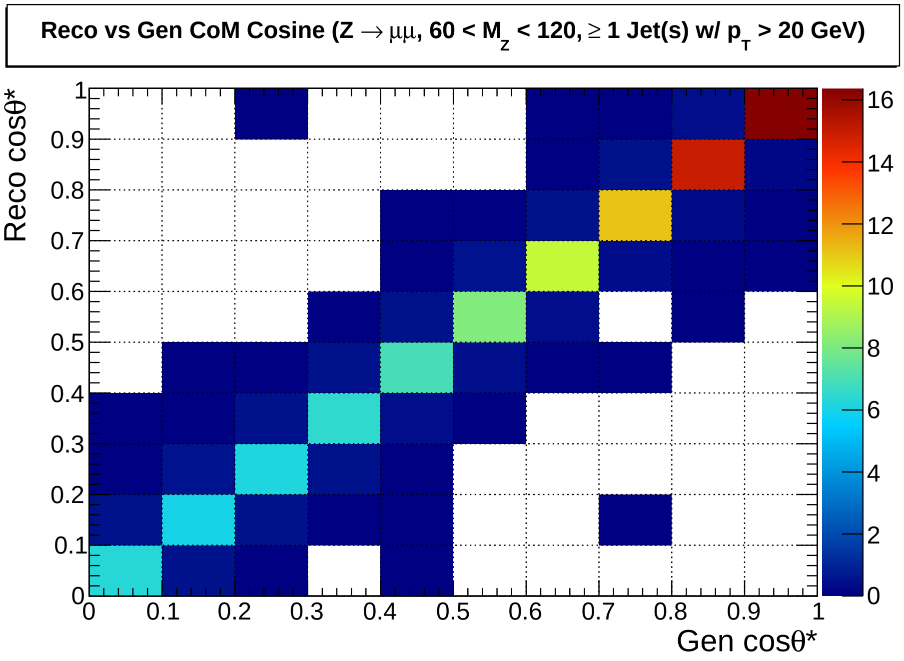


Figure 7.10: Reconstructed *vs.* generated $\cos\theta^*$ in simulated Z+jet signal.

The reconstructed and generated residuals can give us a handle on the uncertainty attributed to detector resolution (see Fig. 7.11). Preliminarily, one can say that the relative uncertainty is better than 2%. However, as mentioned above, unfolding provides a more thorough evaluation and the Bayes unfolding results are shown in Fig 7.12. Unfortunately, we are limited by a lack of collision statistics. Nonetheless, the variation in the mean residual is better than 4% and we use this as a conservative relative uncertainty from resolution effects.

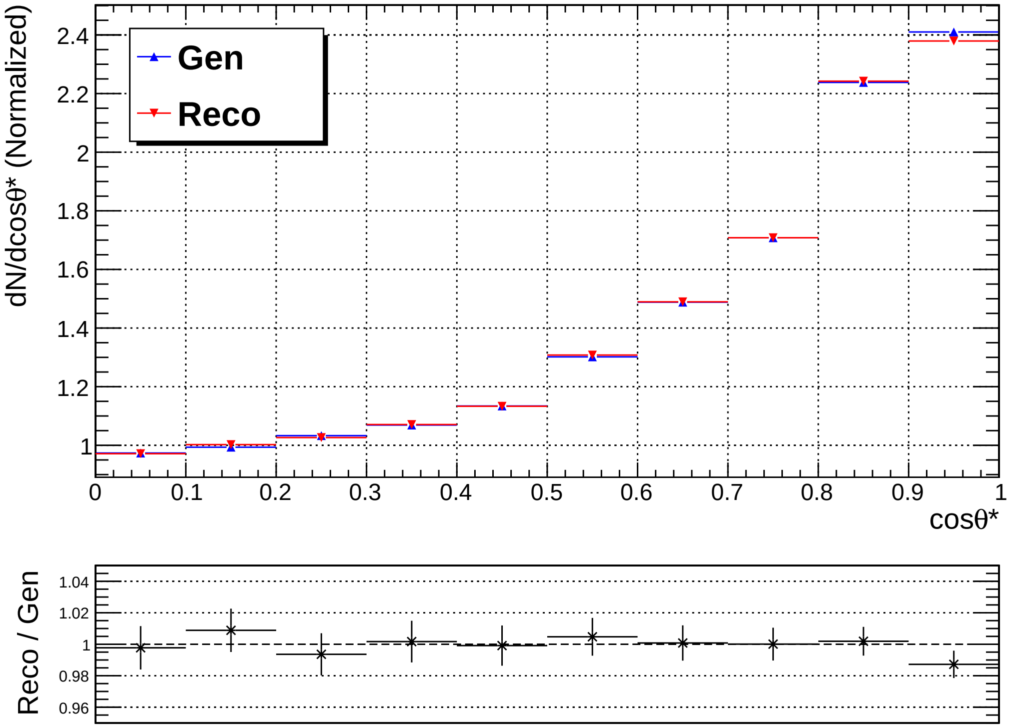


Figure 7.11: Reconstructed and generated $\cos\theta^*$ in simulated Z+jet signal. The bottom plot shows the residuals, i.e., reconstructed divided by generated.

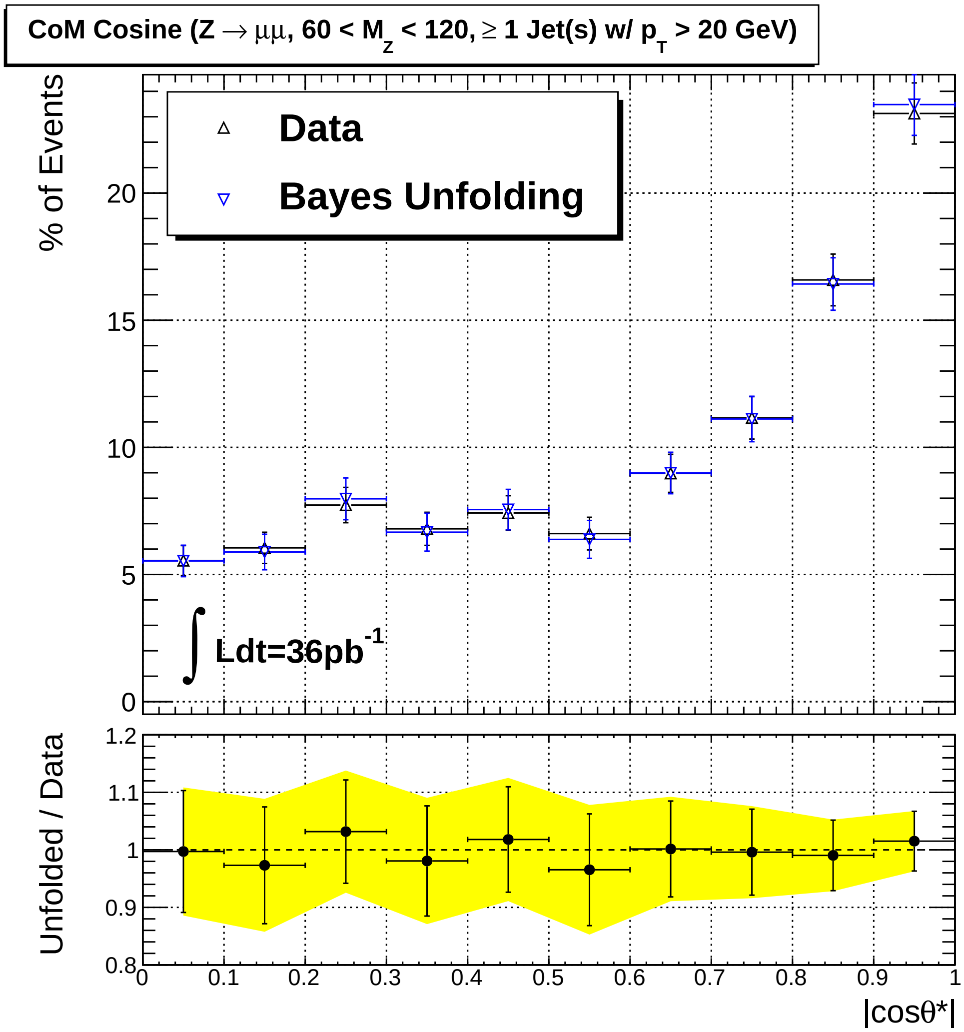


Figure 7.12: Results of the Bayes unfolding technique performed on the measured $\cos\theta^*$ distribution. The bottom plot shows the residuals. The error bar represents the statistical uncertainty while the yellow error band represents a combination of statistical and unfolding uncertainty. The total uncertainty is limited by collision statistics.

Table 7.1: List of relative systematic uncertainties.

Source	Relative Value
Muon energy scale and resolution	$< 1\%$ (not used)
Jet energy scale	$< 5\%$
Jet p_{T} and η resolution (unfolding)	$< 4\%$

CHAPTER 8

RESULTS

The final relative Z^0 angular distribution in the CM frame is shown in Fig. 8.1 (leading-jet) and Fig. 8.2 (multi-jet), after all selections are applied. The normalization is such that the average value for $|\cos\theta^*|$ in the range of 0.0 to 0.3 is unity. Recall that values of $\cos\theta^* > 0.90$ are subject to a phase space bias (see Chapter 7).

Comparing the angular distribution to MC simulation allows us to test the predictions of pQCD, where significant deviations may signify a need for new physics models. For the multi-jet method, the data and simulation comparison has a normalized $\chi^2 = 0.87$ with a probability value [41] of 55.4%. The probability value is a measure of the significance, i.e., how likely is it that the agreement between collision data and simulation is attributed to chance. Comparisons with a probability value above 5% are considered to be sufficiently in agreement [42]. The differences in the single- and multi-jet algorithms are discussed in Appendix C. We conclude that, within statistical and systematic uncertainties, the measured Z^0 angular distribution has no significant deviations from NLO QCD predictions.

Also of interest is the boosted system rapidity for both the leading-jet (Fig. 8.3) and multi-jet (Fig. 8.4) methods. The boosted system rapidity can be factored out of the angular distribution cross section (Eq. 1.7), where y_B only enters in the description of the PDFs. Therefore, comparing the y_B distribution to pQCD predictions allows us to evaluate our knowledge of the PDFs. The difference in deviation between the leading- and multi-jet methods are of interest and will be studied later; these differences may lead to a better modeling of the PDFs. However, systematic studies have not been performed and will be evaluated at a later time. We therefore cannot make any conclusive statements about this distribution yet.

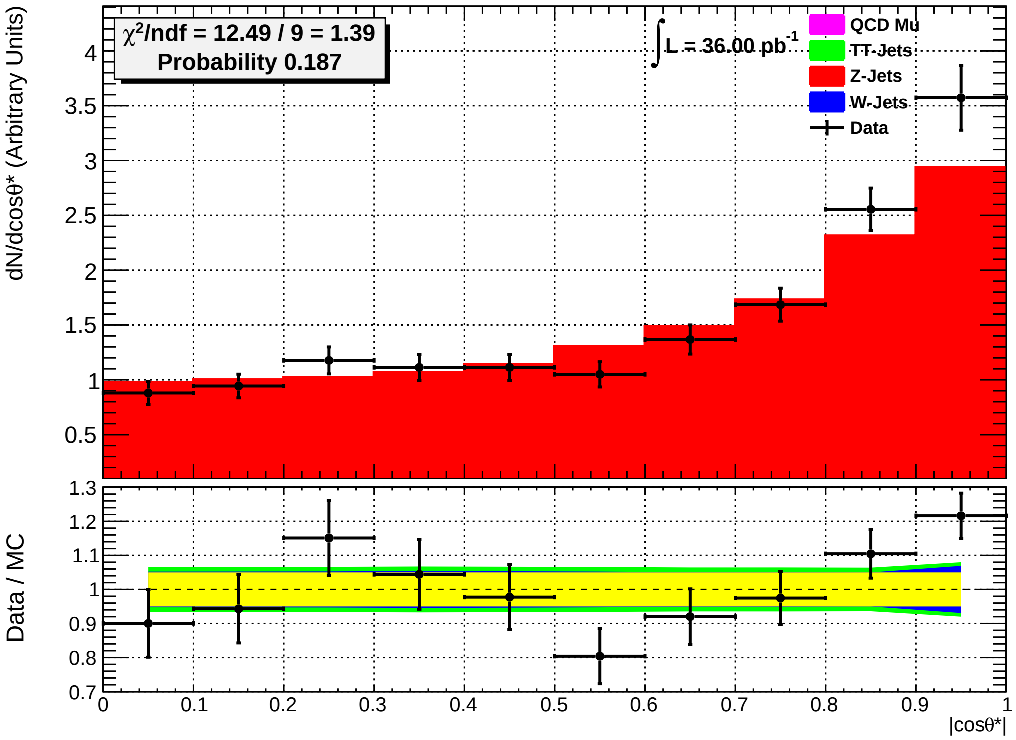


Figure 8.1: The Z^0 CM angular distribution for collision data using the leading-jet method along with simulation residuals. Also shown are the corresponding distributions for signal (red) and various background (other) simulations. For the top plot, the collision data error bars represent a combination of statistical and systematic uncertainties. For the bottom plot, the error bars represent only the statistical uncertainty. The yellow, blue and green bands represent the PDF, PDF + JES, and PDF + JES + Unfolding uncertainties. Values above 0.90 are subject to a phase space bias (see Chapter 7).

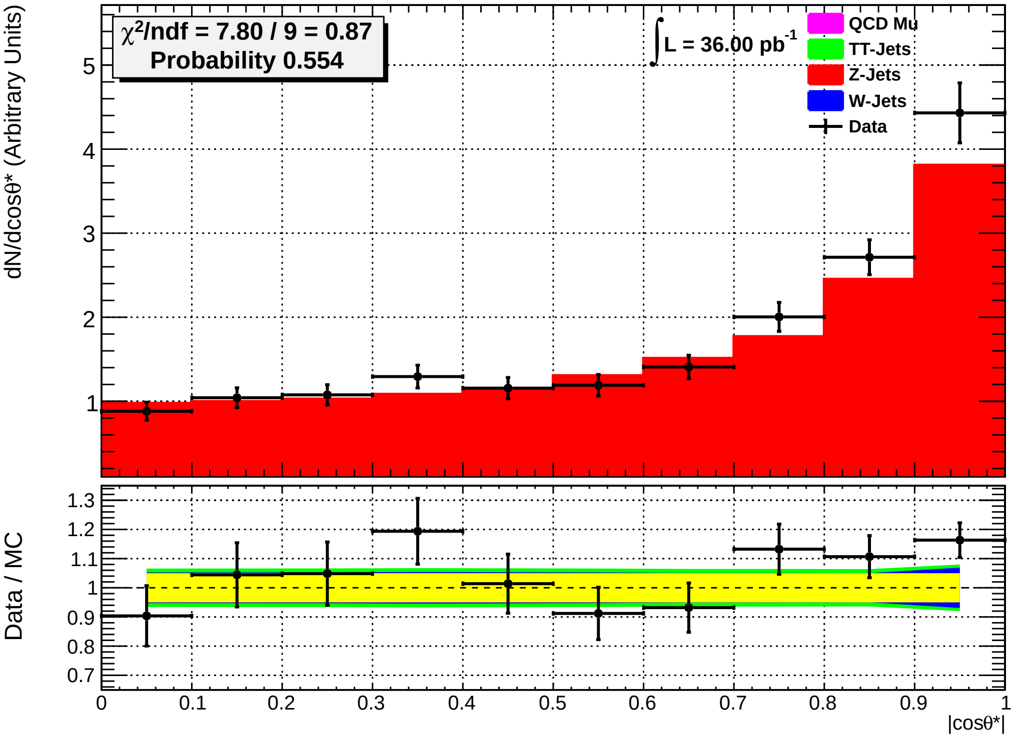


Figure 8.2: The Z^0 CM angular distribution for collision data using the multi-jet method along with simulation residuals. Also shown are the corresponding distributions for signal (red) and various background (other) simulations. For the top plot, the collision data error bars are a combination of statistical and systematic uncertainties. For the bottom plot, the error bars represent only the statistical uncertainty. The yellow, blue and green bands represent the PDF, PDF + JES, and PDF + JES + Unfolding uncertainties. Values above 0.90 are subject to a phase space bias (see Chapter 7).

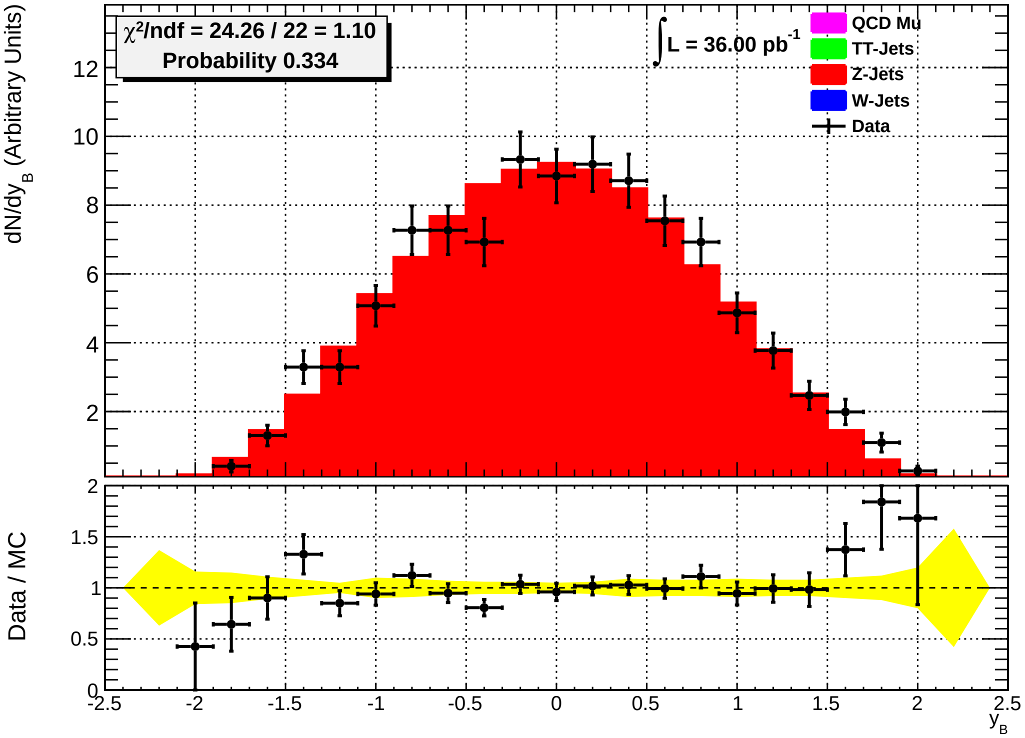


Figure 8.3: The boosted system rapidity distribution for collision data using the leading-jet method, along with simulation residuals. Also shown are the corresponding distributions for signal (red) and various background (other) simulations. The error bars represent only the statistical uncertainty. The yellow bands represents the PDF uncertainty.

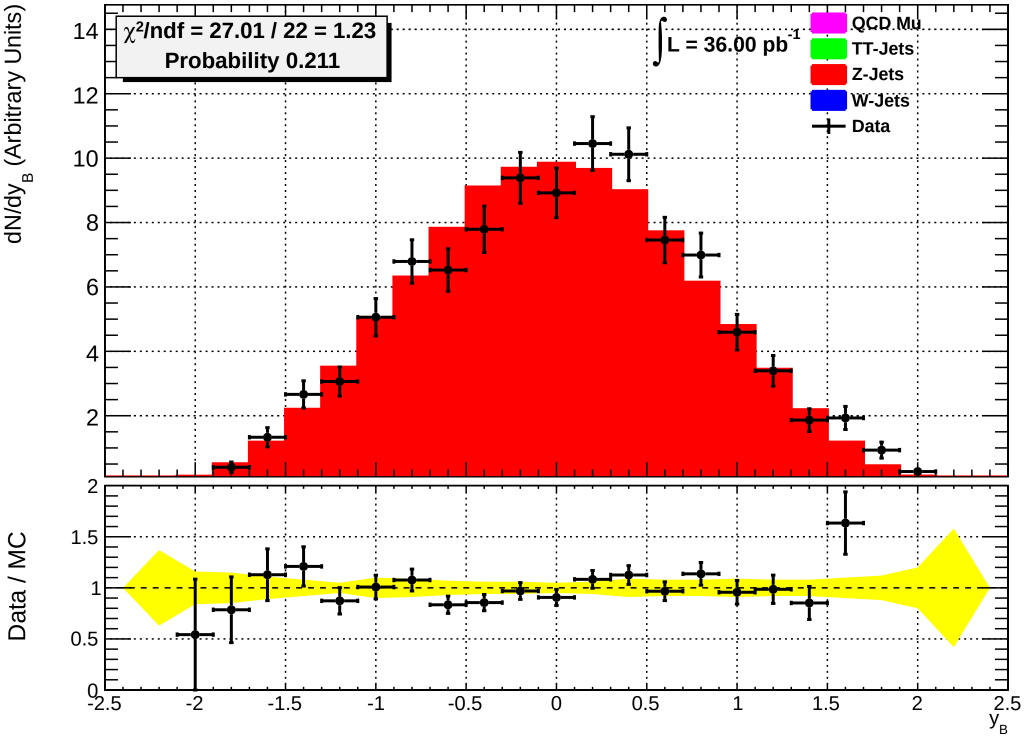


Figure 8.4: The boosted system rapidity distribution for collision data using the multi-jet method, along with simulation residuals. Also shown are the corresponding distributions for signal (red) and various background (other) simulations. The error bars represent only the statistical uncertainty. The yellow bands represents the PDF uncertainty.

8.1. Conclusions

For the first time ever, the Z^0 ($\rightarrow \mu\mu$) boson angular distribution in the parton-parton CM frame was measured at $\sqrt{s} = 7$ TeV for Z+jet events. The data sample corresponds to an integrated luminosity of approximately 36 pb^{-1}. Events in which there is a Z^0 and at least one jet, with a jet p_T threshold of 20 GeV and jet $|\eta|$ less than 2.4, were selected for the analysis. Event reconstruction was performed using particle flow, an algorithm designed to optimizing particle identification by exploiting information from all subdetectors. Events were corrected for pile-up and were unfolded to correct for bin migrations attributed to detection efficiency and measurement resolutions. The CM phase space was flattened to correct for biases introduced by lab frame selections. The measurement was compared to the MadGraph MC generator, which simulates the latest pQCD predictions. Within experimental and theoretical uncertainties, the measured angular distribution is in agreement with NLO pQCD predictions; showing that the SM is valid at the LHC energy regime.

Outlook

As the LHC continues to collect data, the measurement of the Z+jet angular distribution can become more precise and farther reaches of $\cos\theta^*$ may be reached; allowing for higher precision tests of the SM at LHC energies. Higher precision measurements can lead to better modeling of pQCD predictions, improved calibrations of the detector and tighter constraints on the Higgs background. As of Summer 2011, the LHC accelerator was delivering in a single 12-hour fill the same amount of data that was delivered in the whole 2010 period. With the current integrated luminosity, there should be sufficient statistics to extend the measurement while tightening the selection criteria utilized to correct for the phase space bias. Furthermore, the systematic uncertainties currently limited by statistics will diminish. The CMS detector has recorded a total of over 5 fb^{-1} and will continue to record proton collision events until the end of the 2012 run. From a theoretical point-of-view, future NNLO generators with $2 \rightarrow 3, 4, 5, ...$ processes will also allow more detailed comparisons to pQCD and allow us to constrain PDF models.

REFERENCES

[1] Donald H. Perkins. *Introduction to High Energy Physics*. Cambridge University Press, 4th edition, 2000.

[2] Dan Green. *High p_{T} Physics at Hadron Collider*. Number 22 in Cambridge monographs on particle physics, nuclear physics, and cosmology. Cambridge University Press, 2005.

[3] Sidney D. Drell and Tung-Mow Yan. Massive lepton-pair production in hadron-hadron collisions at high energies. *Phys. Rev. Lett.*, 25:902–902, Sep 1970.

[4] R. K. Ellis, W. J. Stirling, and B. R. Webber. *QCD and Collider Physics*, volume 8 of *Cambridge Monographs on Part. Phys., Nucl. Phys. and Cosm.* Cambridge, 1996.

[5] CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[6] CDF Collaboration. W boson + jet angular distribution in p$\bar{p}$ collsions at $\sqrt{s} =$ 1.8 TeV. *Phys. Rev. Let.*, 73:2296–2300, 1994.

[7] Leslie F. Nakae. Direct photon center-of-mass angular distributions in proton - anti-proton collisions at S**(1/2)=1.8-TeV. *FERMILAB-THESIS-1992-19*, 1992. Ph.D. Thesis (Advisor: James Bensinger).

[8] Paul Michael Rubinov. The Center-of-mass angular distribution of direct photons at S**(1/2) = 1.8-TeV observed with the D0 detector. *FERMILAB-THESIS-1995-72*, 1995. Ph.D. Thesis (Advisor: Guido Finocchario).

[9] V. et al. Khachatryan. Measurement of dijet angular distributions and search for quark compositeness in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev. Lett.*, 106:201804, May 2011.

[10] ATLAS Collaboration. High-pt dijet angular distributions in pp interactions at $\sqrt{s}$ =7 tev measured with the atlas detector at the lhc. Technical Report ATLAS-CONF-2010-074, CERN, Geneva, Jul 2010.

[11] CDF Collaboration. W Boson + Jet Angular Distribution in $p\bar{p}$ Collsions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev. Lett.*, 73:2296–2300, Oct 1994.

[12] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*. CERN, Geneva, 2004.

[13] CMS Collaboration. Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS. *arXiv e-prints*, 2011.

[14] CMS Collaboration. Cms computing: Technical design report. *Technical Design Report CMS*, 7(023), June 2005.

[15] CMS Collaboration. Particle-flow event reconstruction in CMS and performance for jets, taus, and MET. *CMS Physics Analysis Summary*, CMS-PAS-PFT-09-001, 2009.

[16] CMS Collaboration. Commissioning of the particle-flow reconstruction in minimum-bias and jet events from pp collisions at 7 TeV. *CMS Physics Analysis Summary*, CMS-PAS-PFT-10-002, 2010.

[17] CMS Collaboration. Performance of muon identification in pp collisions at $\sqrt{s} =$ 7 TeV. *CMS Physics Analysis Summary*, CMS-PAS-MUO-10-002, 2010.

[18] Gavin Salam. Towards jetography. *The European Physical Journal C - Particles and Fields*, 67:637–686, 2010. 10.1140/epjc/s10052-010-1314-6.

[19] M. Cacciari, G. P. Salam, and G. Soyez. The anti-kt jet clustering algorithm. *JHEP*, 4:63, 2008.

[20] F. Maltoni and T. Stelzer. MadEvent: Automatic event generation with MadGraph. *JHEP*, 2:027, 2003.

[21] T. Sjostrand, S. Mrenna, and P. Skands. PYTHIA 6.4 physics and manual. *JHEP*, 5:026, 2006.

[22] R. Field. Early LHC underlying event data - findings and surprises. Invited talk at HCP2010, 2010.

[23] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002.

[24] GEANT4 Collaboration. GEANT4 - a simulation toolkit. *Nucl. Instrum. Meth.*, A506(3):250–303, 2003.

[25] J. Allison et al. GEANT4 developments and applications. *IEEE Trans. Nucl. Sci*, 53(1):270–278, 2006.

[26] K. Melnikov and F. Petriello. Electroweak gauge boson production at hadron colliders through $o\left(\alpha_s^2\right)$. *Phys. Rev. D*, 74(11):114017, 2006.

[27] R. Kleiss and W. J. Stirling. Top quark production at hadron colliders: Some useful formulae. *Z. Phys.*, 40(3):419–423, 1988.

[28] CMS Collaboration. Tracking and primary vertex results in first 7 TeV collisions. *CMS Physics Analysis Summary*, CMS-PAS-TRK-10-005, 2010.

[29] CMS Collaboration. Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at sqrt(s) = 7 TeV. *arXiv e-prints*, 2011.

[30] CMS Collaboration. Jet Production Rates in Association with W and Z Bosons in pp Collisions at sqrt(s) = 7 TeV. 2011. Submitted to JHEP.

[31] A. V. Lipatov, M. A. Malyshev, and N. P. Zotov. Testing for k_T-factorization with inclusive prompt photon production at LHC. *Physics Letters B*, 699:93–97, May 2011.

[32] S. Alekhin et al. The PDF4LHC Working Group Interim Report. *ArXiv e-prints*, January 2011.

[33] Hung-Liang Lai et al. New parton distributions for collider physics. *Phys. Rev.*, D82:074024, 2010.

[34] H. Baer, J. Ohnemus, and J. F. Owens. Next-to-leading-logarithm calculation of direct photon production. *Phys. Rev. D*, 42:61–71, Jul 1990.

[35] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009.

[36] G. Cowens. A survey of unfolding methods for particle physics. In *Proc. Advanced Statistical Techniques in Particle Physics*, Durham, 2002.

[37] T. Adye. Unfolding algorithms and tests using RooUnfold. *ArXiv e-prints*, May 2011.

[38] G. D'Agostini. A Multidimensional unfolding method based on Bayes' theorem. *Nucl. Instrum. Meth.*, A362:487–498, 1995.

[39] Andreas Hocker and Vakhtang Kartvelishvili. SVD Approach to Data Unfolding. *Nucl. Instrum. Meth.*, A372:469–481, 1996.

[40] P. C. Hansen. The l-curve and its use in the numerical treatment of inverse problems. In *in Computational Inverse Problems in Electrocardiology, ed. P. Johnston, Advances in Computational Bioengineering*, pages 119–142. WIT Press, 2000.

[41] N. D. Gagunashvili. Comparison of weighted and unweighted histograms. *ArXiv Physics e-prints*, May 2006.

[42] Thomas Sellke, M. J Bayarri, and James O Berger. Calibration of ρ Values for Testing Precise Null Hypotheses. *The American Statistician*, 55(1):62–71, 2001.

APPENDIX A

UNFOLDING

Unfolding was discussed in Chapter 7, where we chose to perform a Bayes technique (see Fig. 7.12). However, we use the bin-by-bin (Fig. A.1) and matrix inversion (Fig. A.2) methods as cross-checks. We find that all unfolding techniques are in agreement and have better than 4% relative uncertainties (we use the variation in the mean since we are limited by statistics).

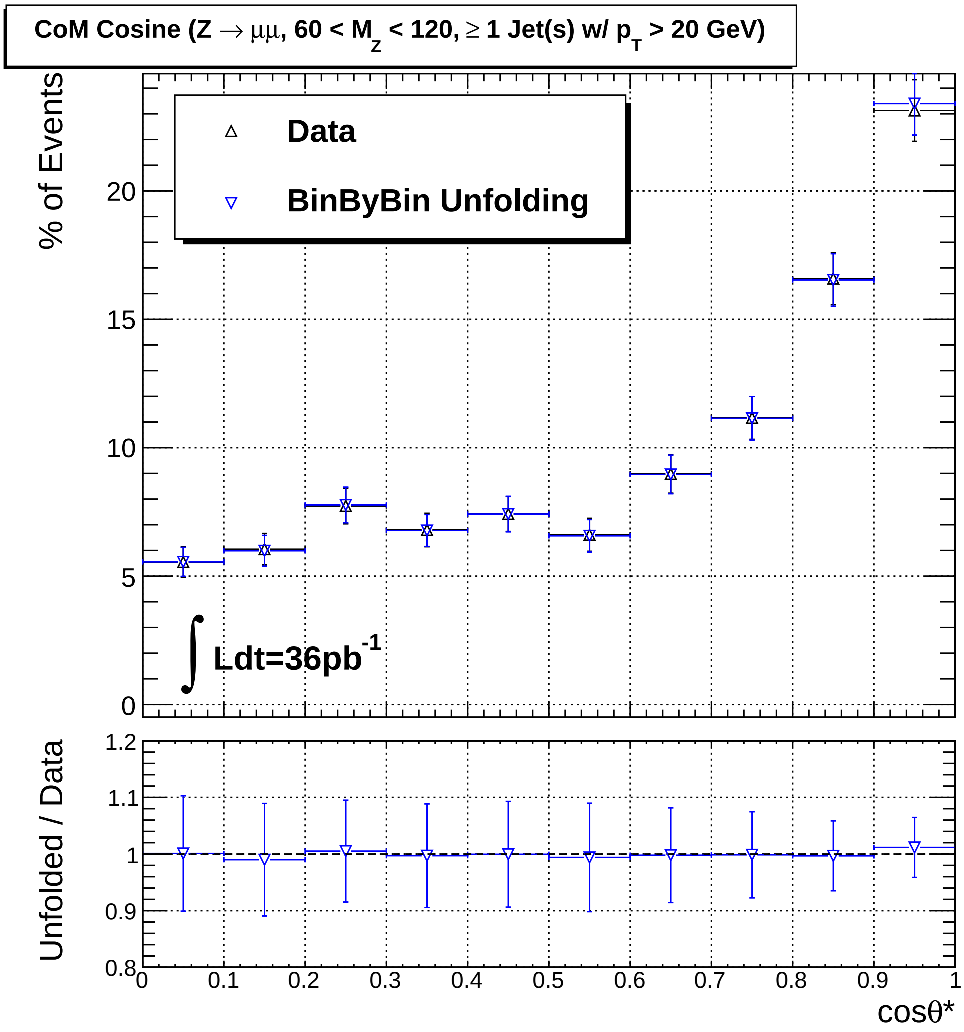


Figure A.1: Results of the bin-by-bin unfolding technique performed on the measured $\cos\theta^*$ distribution. The bottom plot shows the residuals, where the uncertainties are limited by statistics.

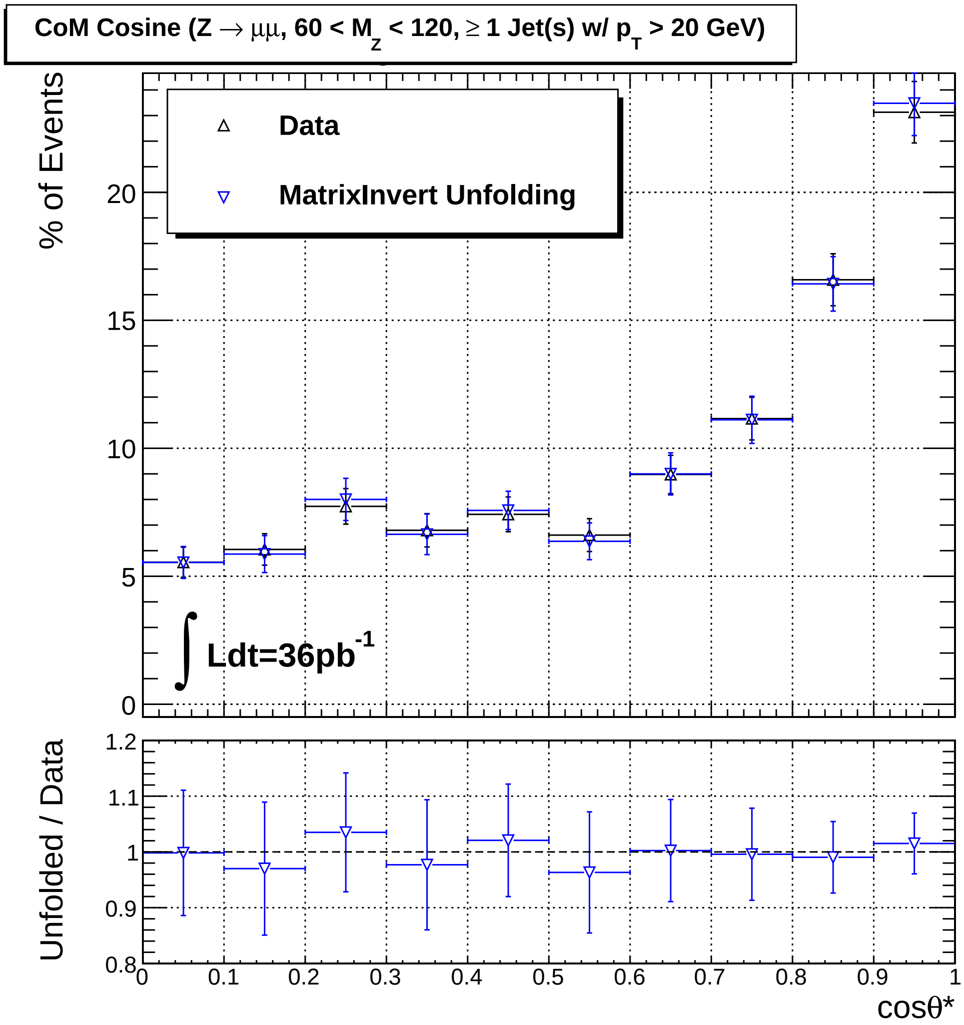


Figure A.2: Results of the matrix inversion unfolding technique performed on the measured $\cos\theta^*$ distribution. The bottom plot shows the residuals, where the uncertainties are limited by statistics.

APPENDIX B

MUON ETA EFFICIENCY

A muon selection dependence on η was discussed in Chapter 7 and shown in Fig. 7.1. This effect was studied in two ways: by correcting for the efficiency using a parameterization and by applying a bin-by-bin correction. Figure B.1 shows the efficiency parameterization,

$$\epsilon = p_0 + p_1\eta + p_2\eta^2, \tag{B.1}$$

where p_0, p_1, and p_2 are the fit parameters shown in the figure. This parameterization flattens the relatively large inefficiencies (e.g., $|\eta| \approx 0.1$ and 1.7) and seeks to correct for the general η dependence; the correction is negligible and shown in Fig. B.2. We also performed a bin-by-bin correction to study the effect of the large inefficiencies. Both methods of correction have a $< .05\%$ effect on the angular distribution.

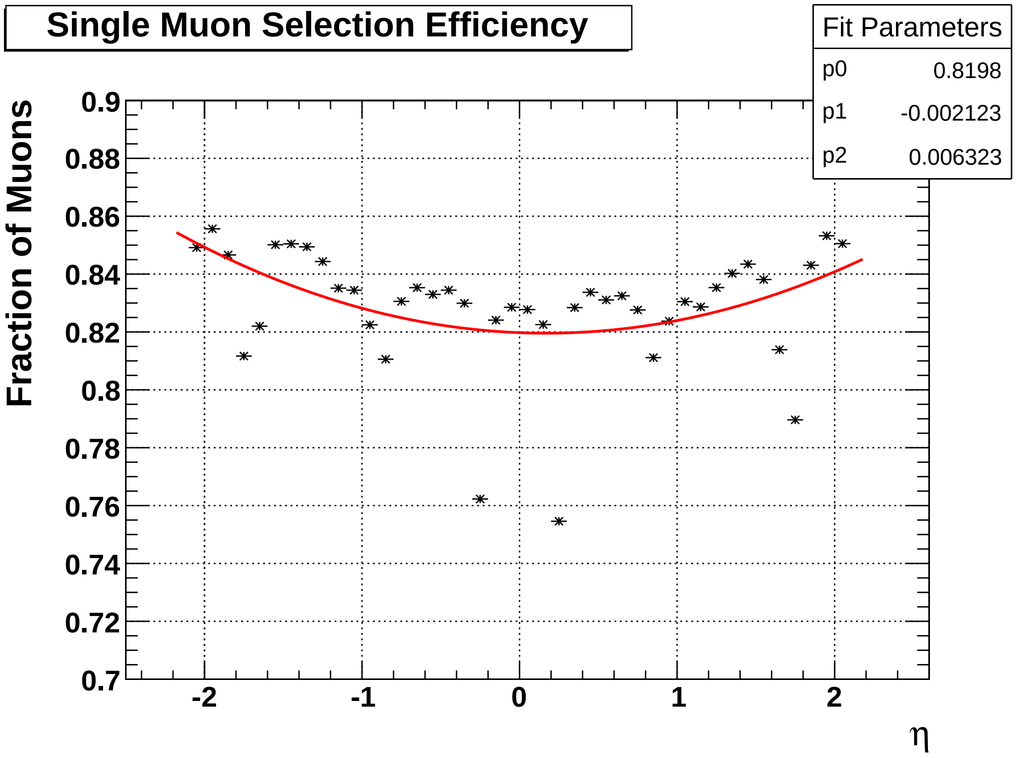


Figure B.1: Fit for the single muon selection efficiency as a function of η for a simulated Z+jet signal. Also shown are the fit parameters.

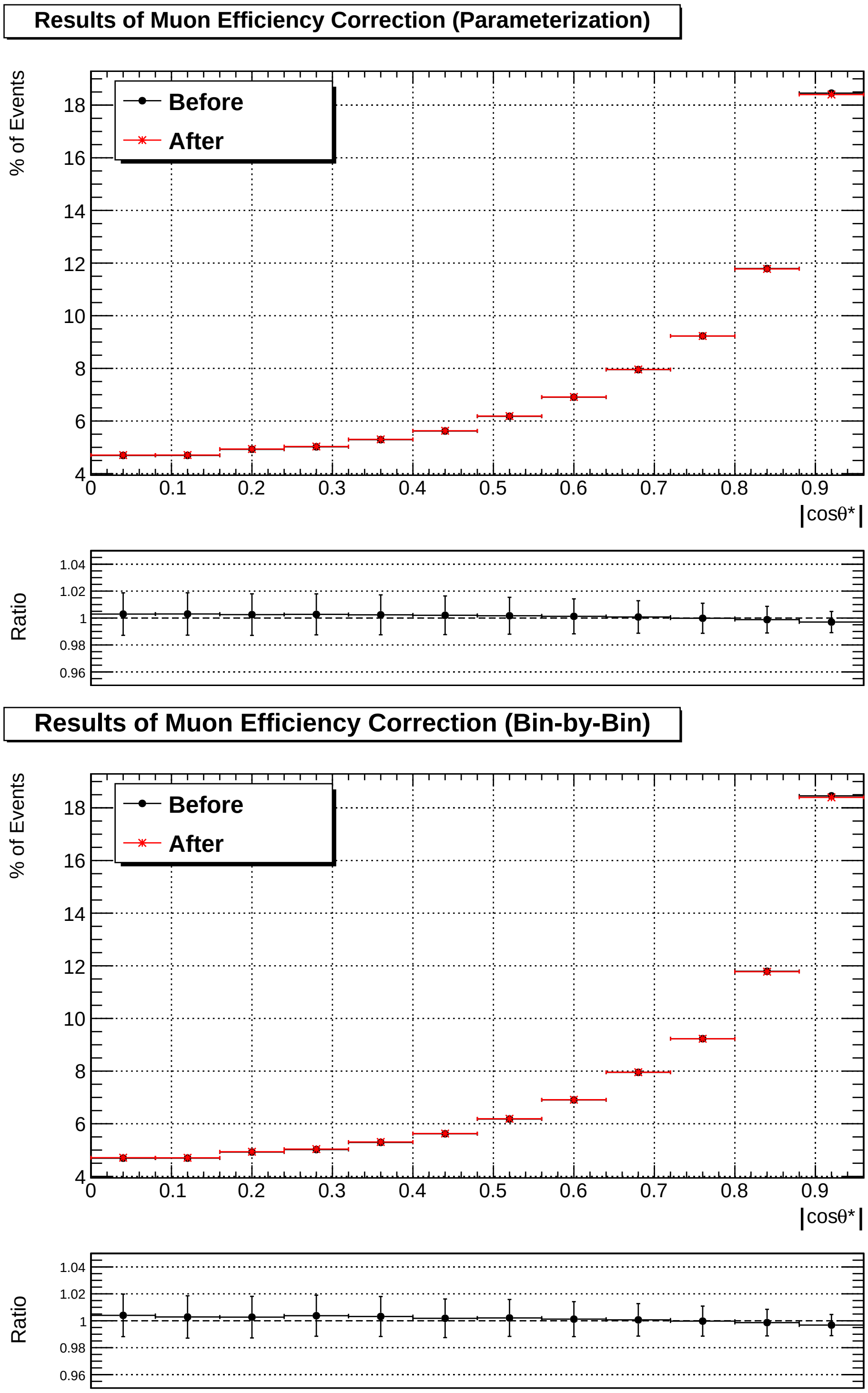


Figure B.2: Results of the muon selection efficiency parameterization (top) and bin-by-bin (bottom) correction, shown for a simulated Z+jet signal. Also shown are the ratios (after/before).

APPENDIX C

SINGLE- AND MULTI-JET DIFFERENCES

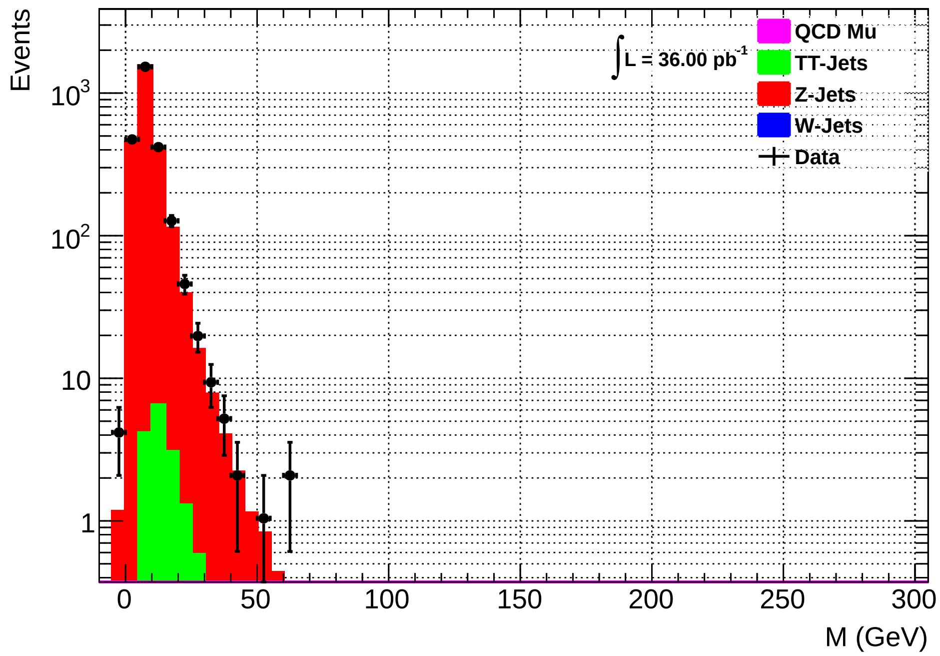
As mentioned in Chapter 6, there are two methods of including jets in CM kinematic calculations: the leading-jet method uses the leading p_{T} jet kinematics while the multi-jet method uses the sum of all jet four-momentum vectors. There is no motivation for choosing between the leading- and multi-jet methods; the multi-jet method intuitively accounts for radiation as it would recombine radiated jets with their parent, but it does not make sense to combine multiple hard-scatter jets (i.e., partons from the ME). Below are the noticeable differences in Z^0 and jet distributions for the leading- and multi-jet algorithms.

- Jet mass (Fig. C.1): the mass of a multi-jet will generally be greater than that of a single jet since the multi-jet method combines multiple jets that may be well separated in η (the mass of a composite jet can be approximated as $M \sim E_1 E_2 \theta$, where E_1 and E_2 are the energy deposits in calorimeter towers and θ is the polar angle between them).

- Jet p_{T} (Fig. C.2): although we apply a jet $p_{\mathrm{T}} > 20$ GeV selection on individual jets, the multi-jet method vectorially sums the four-momentum of jets and may result in a jet with $p_{\mathrm{T}} \leq 20$ GeV

As for the angular distributions (Figs. 8.1 and 8.2), the multi-jet has a better agreement with NLO QCD because it better describes the event; the leading-jet method ignores radiation effects.

Leading-Jet



Multi-Jet

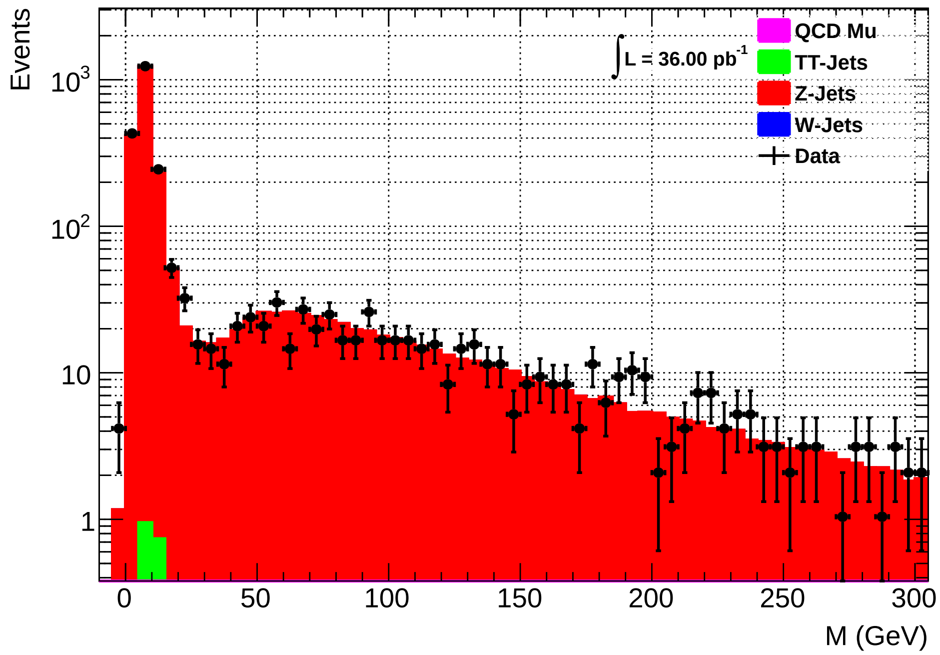
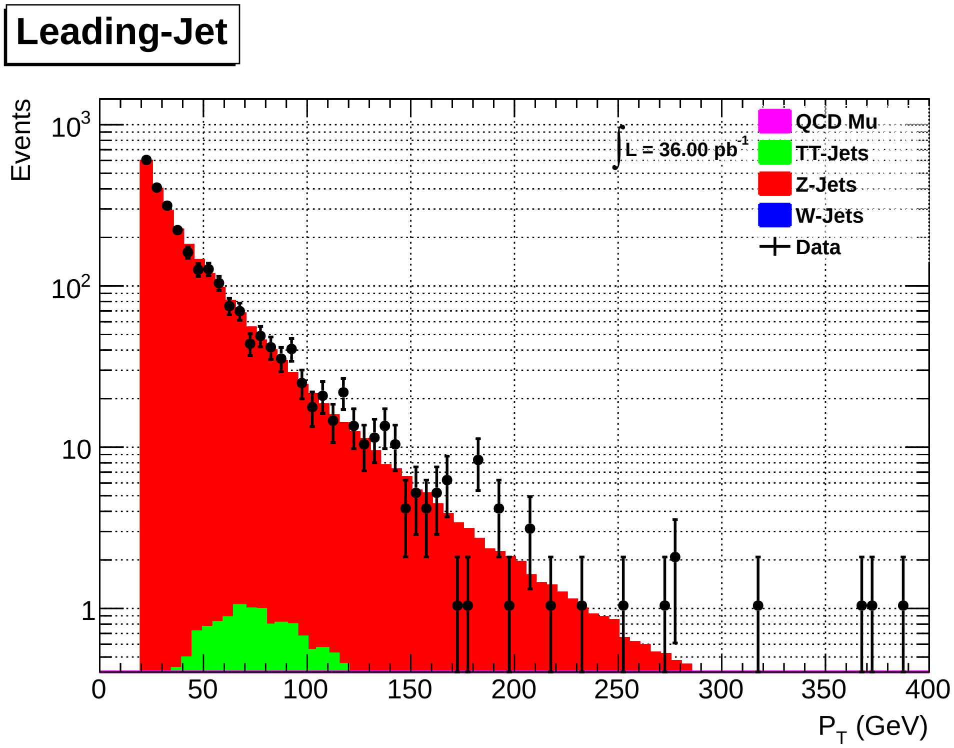


Figure C.1: Jet mass for the leading- (top) and multi-jet (bottom) methods. Collision data (marker) and signal simulations (red) are shown. Background simulations are also shown, but are negligible.



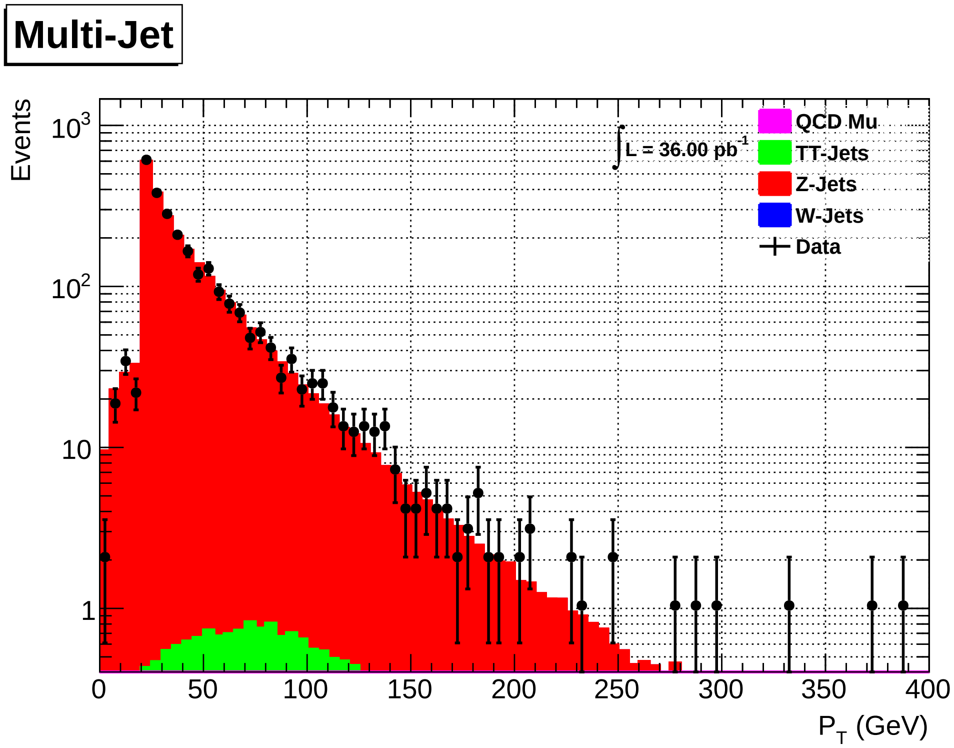


Figure C.2: Jet p_{T} for the leading- (top) and multi-jet (bottom) methods. Collision data (marker) and signal simulations (red) are shown. Background simulations are also shown, but are negligible.

APPENDIX D

CMSSW PHYSICS OBJECTS

This appendix lists the CMSSW 4_2_5 CVS branch "tags" used for the analysis (Table D.1), along with select methods available in the CMSSW physics analysis toolkit (PAT) muon and jet classes.

Table D.1: CMSSW 4_2_5 CVS tags used for the analysis.

Package	Tag
CommonTools/RecoAlgos	V00-03-13
DataFormats/PatCandidates	V06-04-18
FWCore/GuiBrowsers	V00-00-57
PhysicsTools/PatAlgos	V08-06-38
PhysicsTools/PatExamples	V00-05-22
PhysicsTools/SelectorUtils	V00-03-17
PhysicsTools/Utilities	V08-03-09
RecoJets/Configuration	V02-04-17
RecoTauTag/Configuration	V01-02-02
RecoTauTag/RecoTau	V01-02-06
RecoTauTag/TauTagTools	V01-02-00

PAT Muon Class

The list of methods was adapted from the CMS Doxygen reference manual and are given with minimal explanation.

charge() particle electric charge

energy() particle energy

et() particle transverse energy

eta() particle pseudorapidity

mass() particle mass

mt() particle transverse mass

p() particle momentum
p4() particle four-momentum
phi() particle azimuth
pt() particle transverse momentum
pdgId() particle PDG ID
rapidity() particle rapidity

dB() track impact parameter
globalTrack() track reconstructed in both tracker and muon detector
normChi2() normalized chi-square of the global track
numberOfValidHits() returns the number of valid hits on the global track
vertex() track vertex position

ecalIso() ECAL isolation energy
hcalIso() HCAL isolation energy
trackIso() tracker isolation energy

resolE() resolution of energy
resolEta() resolution of eta
resolP() resolution of momentum
resolPhi() resolution of azimuth
resolPt() resolution of transverse momentum

genParticle() the matched generator particle

PAT Jet Class

The PAT jet class has methods similar to the Muon class, except for the following:

associatedTracks() tracks associated with the jet
correctedJet(...) corrected jet based on input level
jetArea() area ($\eta \times \phi$) of jet
jetCharge() charge of jet
pileup() calculated pileup energy contribution

chargedEmEnergyFraction() fraction of charged electromagnetic jet energy
chargedHadronEnergyFraction() fraction of charged hadronic jet energy
neutralEmEnergyFraction() fraction of neutral electromagnetic jet energy
neutralHadronEnergyFraction() fraction of neutral hadronic jet energy
muonEnergyFraction() fraction of muonic jet energy
photonEnergyFraction() fraction of photonic jet energy

chargedHadronMultiplicity() number of charged hadronic particles in the jet
chargedMultiplicity() number of charged particles in the jet
electronMultiplicity() number of electrons in the jet
muonMultiplicity() number of muons in the jet
neutralHadronMultiplicity() number of neutral hadrons in the jet
neutralMultiplicity() number of neutral particles in the jet
photonMultiplicity() number of photons in the jet

genJet() the matched generated jet
genParton() the matched generated parton

VITA

LUIS LEBOLO

November 22, 1982	Born, Hialeah, Florida
2005	Ronald E. McNair Post-Baccalaureate Achievement Program Florida International University Miami, Florida
	B.S., Physics Florida International University Miami, Florida
2009	M.S., Physics Florida International University Miami, Florida

PUBLICATIONS AND PRESENTATIONS

CMS Collaboration, “Jet Production Rates in Association with W and Z Bosons in pp Collisions at sqrt(s) = 7 TeV”, Submitted to JHEP (2011). arXiv:1110.3226v1

CMS Collaboration, “Measurement of the Inclusive Jet Cross Section in pp Collisions at sqrt(s) = 7 TeV”, Submitted to PRL (2011). arXiv:1106.0208v1

CMS Collaboration, “Measurements of Inclusive W and Z Cross Sections in pp Collisions at sqrt(s) = 7 TeV”, JHEP 01 (2011) 080. doi:10.1007/JHEP01(2011)080.

CMS Collaboration, “Performance of CMS Hadron Calorimeter Timing and Synchronization Using Test Beam, Cosmic Ray, and LHC Beam Data”, JINST 5 (2010) T03013. doi:10.1088/1748-0221/5/03/T03013.

A. Heering et al., “Large-Area SiPMs for the CMS Hadron Outer Calorimeter”, IEEE NSSCR 2 (2007) 1545. doi:10.1109/NSSMIC.2007.4437293.

Analysis of the Laser Calibration System for the CMS HCAL at CERN’s LHC, Presented at the SESAPS’05 Meeting of the American Physical Society (2005)

Z+Jet Center of Momentum Angular Distribution using the Compact Muon Solenoid, Presented at the SESAPS’11 Meeting of the American Physical Society (2011)